DOCUMENT RESUME

ED 095 205 TM 003 877

TITLE	Summary of Research Projects 1972-1973.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	R-53
PUB DATE	Sep 73
NOTE	199p.; For related document, see ED 056 076
EDRS PRICE	MF-$0.75 HC-$9.00 PLUS POSTAGE
DESCRIPTORS	Cognitive Development; Computer Science; Creativity Research; Cultural Differences; Early Childhood Education; *Educational Research; Higher Education; *Human Development; Learning; Personality; Primary Education; *Psychological Studies; Psychometrics; *Research Methodology; Research Projects; *Research Reviews (Publications); Secondary Education; Social Behavior; Testing; Vocational Education

ABSTRACT

Purpose, progress, and in most cases, major findings are presented for research conducted at Educational Testing Service, Princeton, N.J. during 1972-73. Under the rubric, Psychological Research, summaries are provided for work dealing with human development, learning and cognition, personality, and social behavior. For the category, Educational Applications, summaries are presented for early and primary education, secondary education, higher education, graduate and professional education, cultural diversity and the disadvantaged, career and vocation, and evaluation of educational systems. In the third and final section, Theories, Methods, and Systems Research, summaries are given for psychometric theory, methodological research, and computer studies. The authors for all research are listed and questions should be directed to them. If a report on a study is available, it is indicated by a number appearing in the body of the write-up which can be used when requesting a report. (RC)

ED 095205

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SUMMARY OF RESEARCH PROJECTS

1972 - 1973

(Report # 53)

BEST COPY AVAILABLE

Educational Testing Service
Princeton, New Jersey
September 1973

M 003 827

FOREWORD

The "Summary of Research Projects" reports on research carried on at ETS during the period July 1, 1972, through June 30, 1973.

Questions concerning the studies reported herein should be directed to the authors. If a report on a study is available, this will be indicated by a number appearing in the body of the write-up. Please use this number when requesting a report.

TABLE OF CONTENTS

I. PSYCHOLOGICAL RESEARCH

A. Human Development

Baker, A. H.

Developmental Changes in Self/World Differentiation 156-158

Developmental Model of Cognitive Style 124-127

Brooks, J.

Attachment Behavior in Thirteen-Month-Old Opposite Sex Twins (with M. Lewis) 128

Peer Group Interaction in Infancy (with M. Lewis and L. Michalson) 169

Emmerich, W.

Differentiation and Development of Social Norms 136-137

Lewis, M.

Assessment of Infant Socioemotional Development (with L. Michalson) 168

Developmental Study of Attention within the First Two Years of Life (with E. Scott and C. Wilson) 120-122

Evaluation of Infant Intelligence: Infant Intelligence Scores--True or False? (with H. McGurk) 123

Infants' Responses to Pictures of Mother, Self, and Other (with J. Brooks) 129

The Integration of Faces and Voices in Infancy (with L. Rosenwein, C. Wilson, and H. McGurk) 130

Longitudinal Study of Infant Attention and Its Relationship to Subsequent Perceptual-Cognitive Development (with C. Wilson) 131-133

Mothers and Fathers, Girls and Boys: Attachment Behavior in the First Two Years of Life (with M. Weinraub and P. Ban) 134

Mother-Infant Interaction (with S. Painter) 166-167

Lewis, M. (continued)

Self, Other, and Fear: The Reaction of Infants to People (with J. Brooks) 150

Variance and Invariance in the Mother-Infant Interaction: A Cross-Cultural Study (with P. Ban) 135

B. Learning and Cognition

Cahen, L.

Difficulty Factors in Learning to Spell 152

Carroll, J.

Application of the Lognormal Model of Word Frequency Distribution 80-81

Factors in Comprehension of Verbal Material (with R. Freedle) 82-83

International Study of Educational Attainment in French as a Foreign Language 163

Study of Aptitude-Achievement Relationship 78

Freedle, R.

Psycholinguistic Theory and Models 84-85

Hale, G.

Developmental Studies of Children's Incidental Learning and Distractibility 151

Developmental Studies of Selective Attention 170

Harman, H.

Reference Measures for Cognitive and Personality Factors 61

Ward, W.

Creativity, Learning, and Attention 142

Developmental Differences in Stimulus Encoding 153

Individual Difference Parameters in Recognition Memory 154

C. Personality

Damarin, F.

Leveling-Sharpening as a Cognitive Style (with V. McKenna and S. Messick) 67

Frederiksen, N.

Effects of Models on Creative Performance in Relation to Anxiety, Ability, and Sex (with F. Evans) 149

Development of Provisional Criteria for the Study of Scientific Creativity (with W. Ward) 29

Goodenough, D.

The Effect of Stress on Dream Affect (with H. Witkin) 45

Genetic Factors in the Development of Field Dependence and Spatial Abilities: Intercorrelations Among Family Members (with H. Witkin) 46

Genetic Factors in the Development of Field Dependence and Spatial Abilities: X-Linkage Study (with H. Witkin and I. Olkin) 47

Respiratory Correlates of Emotional States (with P. Oltman) 48

Jackson, D.

Convergent and Discriminant Validation of the DPI 88

The Effects of Detoxification on Psychopathology for Alcoholics as Measured by the Differential Personality Inventory 89

A Four-Dimensional Interpretation of Risk Taking 90

Judged Frequency of Endorsement and Frequency of Occurrence Scale Values and Dispersions for MMPI Items (with S. Messick) 91

A Model for Inferential Accuracy 92

Studies in the Factorial Interpretation of Multitrait-Multimethod Matrices in Personality Assessment 93

Kogan, N.

Longitudinal Study of Creativity 141

Messick, S.

Cognition of Affect 143

Cognitive Styles and Attitudinal Consistency 147

Cognitive Styles and Psychopathology 140

Dimensions of Perceived Item Similarity 66

Perceptual Factors in Cognition and Personality (with J. French) 62

Personality Organization in Cognitive and Social Processes 68-74

Oltman, P.

Analytical Ability in Verbal Functioning (with H. Witkin) 49

Psychological Differentiation as Related to Antisocial Behavior, Impulsivity, and Endocrine Functioning in a Prison Population (with H. Witkin and D. Owen) 57

Laterality Study 50

Studies of Young Adults with Uneven Cognitive Development (with H. Witkin) 51

Owen, D.

Relationship of Height to Criteria of Antisocial Behavior . . 58

Stricker, L.

Personality Research Form: Factor Structure and Response Style Involvement 86

Response Styles and 16 PF Higher-Order Factors 87

Retest Changes on Personality Scales 75

Witkin, H.

Cognitive Patterning in Sighted Children with Bilateral Retinoblastoma (with P. Oltman) 52

Cross-cultural Studies of the Development of Psychological Differentiation (with P. Oltman) 54

Cognitive Style as a Factor in Dyadic Interaction Under Conditions of Initial Conflict (with D. Goodenough and P. Oltman) 53

Witkin, H. (continued)

Genetic Factors in Field Dependence and in Aggressive Behavior: Studies in Sex Chromosome Anomalies (with D. Goodenough and D. Owen) 59

Modes of Representation of Presleep Experiences in Dreams (with D. Goodenough) 55

Sex Chromosome Anomalies in a Male Prison Population (with P. Oltman and D. Owen) 60

Teacher-Student Interaction, Student Achievement and Mutual Attraction between Teacher and Students as a Function of Match or Mismatch in Cognitive Style (with D. Goodenough, P. Oltman, C. A. Moore, and F. McDonald) 56

D. Social Behavior

Kogan, N.

Determinants of Risk Taking for Others (with A. Teger) 145

Dimensions and Categories of Physiognomic Judgment 144

Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma (with M. Hermann) . . 76

Motivation and Performance in Adult-Evaluative and Peer-Competitive Contexts 146

Negotiation Behavior Under Conditions of Observation and Consultation 77

II. EDUCATIONAL APPLICATIONS

A. Early and Primary Education

Anderson, S.

CIRCUS 11

Cahen, L.

Manipulated Expectancies and Teacher Bias 155

Chittenden, E.

Assessment and Development of Cognitive Skills in Young Children: A Classroom Approach (with A. Bussis) 177

Study of Teacher Change in Open Educational Settings (with A. Bussis and M. A. Amarel) 178

Emmerich, W.

Preschool Teachers' Beliefs on Effective Teaching Techniques and Their Relationships to Pupil Characteristics 36

Norton, D.

Evaluation of Preschool for Urban Children Project 8

Minneapolis Pyramid Sibling Study, Phase III 9

Powers, D.

Evaluation of the New Approach Method 39

Shipman, V.

Bell Lab Day Care Evaluation 184

Trismen, D.

A Descriptive and Analytic Study of Compensatory Reading Programs 41

B. Secondary Education

Boldt, R.

Accuracy of Self-reported Grades Program 7

Status of Drug-Related Instructional Programs in Secondary and Elementary Educational Institutions (with R. Reilly) 176

Diederich, P.

Newspaper Test 159

Cooperative Literature Tests 79

Evans, F.

SAT Speededness Study 6

Pitcher, B.

College Board Activity and Participation Statistics 2

C. Higher Education

Centra, J.

Departmental Characteristics and Student Academic Growth . . 174

The Effectiveness of Student Feedback in Modifying College Instruction 182

The Relationship between Student and Alumni Ratings of Teachers 32

Self-Ratings of College Teachers: A Comparison to Student Ratings 183

Student Points of View in Ratings of College Instruction . . 179

Evans, F.

GRE-Quantitative Coaching Study 27

Schrader, W.

Quantitative Comparison Validity Study 3

Warren, J.

Assessment of Junior College Student Concerns (Student Reactions to College) 5

Study of College Degrees 165

D. Graduate and Professional Education

Boldt, R.

Factor Analysis of Law School Grades 14

NSF Science Aptitude Trends Study 180

Campbell, J.

A Survey of Skills and Abilities Needed for Graduate Study in Business (with P. Casserly) 44

Evans, F.

Advanced Moderator Variable Study (with D. Rock) 13

Pitcher, B.

An Exploratory Study of the Possibility of Curvilinearity in Law School Validity Study Data 16

Student Subgroups Validity Study 17

Validity Study of Experimental Item Types (with A. Carlson and M. McPeek) 15

Validity Study of LSAT Part-Score Reporting 17

Reilly, R.

Graduate Performance by Critical Incident 23

Schrader, W.

Effect of Differences in College Grading Standards on the Prediction of Law School Grades (with B. Pitcher) . . . 18

Witkin, H.

The Role of Cognitive Style in Students' Academic Evolution from High School Through Graduate and Professional School (with P. Oltman and D. Goodenough) 28

E. Cultural Diversity and the Disadvantaged

Campbell, J.

Test and Job Performance of Various Subgroups within Specified Occupations (with L. Crooks, D. Rock, and M. Mahoney) . . . 181

Word Associations of Black and White College Students 24

Emmerich, W.

Personality Structure and Development in Headstart Children . 35

Feldmesser, R.

Effectiveness of the GRE Fee-Waiver Program 26

Survey of Disparities in Educational Achievements 42

Murphy, R.

Learning System in the Social Sciences 43

Survey of Adult Reading Performance 40

Shipman, V.

ETS-Head Start Longitudinal Study of Disadvantaged Children and Their First School Experiences 37

Stricker, L.

Social Stratification 148

Warren, J.

Chicano Validity Study 5

F. Career and Vocation

Campbell, J.

Development of Interest Measurement Instruments for Assignment of Army Officers to Appropriate Positions (with R. Boldt, R. Flaugher, and D. Rock) 171

Centra, J.

Comparison between Faculty Peer Ratings and Student Ratings 32

Women Doctoral Recipients 25

Echternacht, G.

Development and Validity of an Air Force Vocational Inventory Study (with R. Reilly) 175

Goldstein, L.

The Physical Activities Survey of Police Officers in New Jersey 33

Promotion Examinations for Police Sergeants, Lieutenants and Captains in New Jersey 34

Katz, M.

System of Interactive Guidance and Information 172

Katz, M. (continued)

Validating Effects of SIGI on Career Decision-Making Processes of Students 173

McDonald, F.

Teacher Behavior Research 19

G. Evaluation of Educational Systems

Fremer, J.

National Assessment 30-31

McDonald, F.

A Design for an Accountability System for the New York City School System 12

National Commission on Performance-Based Education 20-21

Norton, D.

Efficiency and Efficacy of Evaluation Activities of the State Division of Vocational and Technical Education 10

Powers, D.

Audit of Project IMPACT (Phase II) 38

Evaluation of the Virginia Beach Extended School Year Program 38

III. THEORIES, METHODS, AND SYSTEMS RESEARCH

A. Psychometric Theory

Gleser, L.

Classification Under the Model of Compound Symmetry 107

A Guide to Probability Theory and Applications (with I. Olkin) 108

Large-Sample Expansions for Null Distributions of Likelihood Criteria (with I. Olkin) 109

Gleser, L. (continued)

Multivariate Statistical Inference Under Marginal Structure. I (with I. Olkin) 110

Multivariate Statistical Inference Under Marginal Structure. II (with I. Olkin) 111

Testing If Two Measuring Procedures Measure the Same Dimension 112

Jennrich, R.

Standard Errors for Rotated Factor Loadings 119

Kristof, W.

A Discussion of Certain Reliability Formulas 99

Factorization of Covariance Matrices in Terms of a Specified Pattern or Structure (with B. Wingersky) 160

On Accuracy in Reliability Estimation 99

On Proportional Profiles in Factor Analysis 138

On a Statistic Arising in Testing Correlation 139

On the Statistics of Reliability Estimation Using Parallel Test Forms 100

Testing a Linear Relation between True Scores of Two Measures 161

Testing Whether a Disattenuated Correlation Is Perfect . . . 162

Lord, F.

Automated Hypothesis Testing (with M. Stocking) 63

The Relative Efficiency of Two Tests 98

Estimating Item Parameters (with M. Wingersky) 64

Strong True-Score Theory (with M. Wingersky and M. Stocking) 65

Rubin, D.

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Random Predictors (with T. Stroud) 114

Stroud, T.

Comparing Regressions When Measurement Error Variances Are Known 115

Stroud, T. (continued)

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors 116

Estimating Variance Components by Combining Unbiased Estimates from Independent Data Sets 117

B. Methodological Research

Boldt, R.

GRE Bayesian Validity Study 22

Browne, M.

Analysis of Covariance Structures 101

Donlon, T.

Differential Weighting of Item Distracters in Data Sufficiency Items 4

Gulliksen, H.

Intercultural Attitude Study (with D. Gulliksen) 94

Learning 95-96

Levine, M.

Conference on Psychophysical Measurement 164

Geometric Interpretations of Magnitude Estimation and Other Direct Psychophysical Measurement Procedures 104

The Generalization Function in Probability Learning Experiments with Infinitely Many Stimuli and Responses . . . 103

Nonadditive Analogues of the Basic Mathematics of Additive Measurement 105

Parameterizing Curves Transformable to the Same Shape (with D. Saxe) 106

Lord, F.

Power Scores Estimated by Item Characteristic Curves 98

SAT Delta Study 97

Rubin, D.

Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study (with T. Stroud) 113

Stroud, T.

Use of INDSCAL in Assessing a Cerebral Deterioration Pattern in Psychiatric Patients 118

C. Computer Studies

Browne, M.

Rotation of Factor Matrices (with M. van Thillo) 102

Lavine, L.

Simulation and Modeling 1

Scheduling Techniques 1

BEST COPY AVAILABLE

Simulation and Modeling (29-42) L. Lavine

PURPOSE: To support planning and early work leading to the design and programming of computerized simulation models that will permit program directors and other managers to evaluate the consequences of alternative policies.

PROGRESS: This project supported four separate simulation efforts this year. These were: (1) Model of Educational Attainment, (2) Shipping Model, (3) Faculty Cost Model, and (4) Higher Education Admissions Model. The Model of Educational Attainment was programmed by two summer employees; presentations were made to Messrs. Turnbull, R. Smith and Schabacker. Planning was completed for a shipping model to be tested in the ATSGSB program; further work will be supported by program funds. Early planning was accomplished for the Faculty Cost Model that was later supported and used by IED in their work with a consortium of 12 Pennsylvania colleges. Preliminary work on a Higher Education Admissions Model was completed with G. Schlekat; further work will depend upon support from other areas.

Scheduling Techniques (29-43) L. Lavine

PURPOSE: Research, development and adaptation of existing programs concerned with general problems of planning and scheduling of time and cost.

PROGRESS: PMS (IBM's Project Management System) has been successfully used in the preparation of at least 24 proposals in the course of the year. The basic principles involved in using PMS were presented by Andy Wagner as part of a seminar on proposal preparation. Acceptance of PMS as a generally useful tool is spreading as evidenced by its use in scheduling and controlling the Flexible Benefits project, by its use by Tom Dwyer for Deadline Date Scheduling, and by a request for a course which will be presented by Andy Wagner in late 1973. Work has begun on evaluation of a CALCOMP program, AUTONET, that makes use of the CALCOMP digital plotter to draw a PERT chart based upon scheduling information processed by PMS. ETS's use of PMS for PERT scheduling was described in a paper presented by Andy Wagner at the 1973 meetings of the American Educational Research Association in New Orleans.

College Board Activity and Participation Statistics (CEEB, 130-2) — B. Pitcher

PURPOSE: This study is concerned with the preparation of selected descriptive statistics on College Board candidates to describe patterns of test participation with particular reference to shifts that may affect score interpretation and program planning.

PROGRESS: Candidates and Tests, 1971-72 was completed and distributed.

MAJOR FINDINGS: Candidate and test volume in the admissions testing program continued to follow a downward trend in 1971-72. Four indices of overall testing activity used in the admissions testing program showed the following changes in 1971-72 over 1970-71 volume: (1) a decrease of 2.9% in number of Scholastic Aptitude Tests administered, (2) a decrease of 14.6% in number of Achievement Test series, (3) a decrease of 5.8% in total test volume, and (4) a decrease of 4.7% in number of registrations. The number of Preliminary Scholastic Aptitude Tests increased 2.4%.

Quantitative Comparison Validity Study
(CEEB, 130-22)

W. Schrader

PURPOSE: To obtain evidence on the relative effectiveness of a 30-minute Quantitative Comparison Test and of SAT-Mathematical for predicting college grades.

PROGRESS: A report has been written and was issued as a Statistical Report (SR-73-60).

MAJOR FINDINGS: Validity coefficients for 16 student groups in 12 colleges showed no clear pattern of superiority for either test. There was some tendency, however, for differences in validity between the two tests to be associated with score level on SAT-mathematical. Quantitative Comparison tended to yield higher validities than SAT-mathematical in relatively high-scoring groups; in low-scoring groups, the reverse tendency held. In five of the 16 comparisons, differences between validity coefficients for the two tests were significant at the 5% level.

Differential Weighting of Item Distracters in Data Sufficiency Items (CEEB, 141-5) — T. Donlon

PURPOSE: Differential weighting of wrong-answer responses to multiple-choice items may provide additional information not currently reflected in the conventional scoring methods. Data sufficiency items would seem to offer particular promise in that the meaning of responses is constant for all items regardless of content. A comparison will be made of the efficiency of an a priori weighting system and a system using the judgments of mathematics specialists.

PROGRESS: The experimental tests were prepared and administered to approximately 1800 freshman students at seven colleges. Judged estimates of option merits have been secured, and grade-point averages have been received from the schools. Preliminary analyses of the data revealed that none of a number of empirically or rationally derived composite scores, based upon different linear combinations of the different response category scores, yielded an alternate forms reliability as high as the Total Rights score alone.

A number of additional analyses are being performed. They include determining the composite weights by canonical correlational analysis, principal components analysis, and multiple regression analysis, using the Mathematics Achievement Test as the criterion. Validities are being determined for those linear combinations that have a reliability greater than the Total Rights score reliability. Related research in this area by Hendrickson, Reilly and Jackson, and Davis is being reviewed to adapt the analyses most appropriately.

Assessment of Junior College Student Concerns (Student Reactions to College) (CEEB, 141-10) — J. Warren

PURPOSE: An instrument is being developed to assess the views of junior college students about their college experiences and the congruence or disjunction they see between their experiences and their reasons for being in college. The instrument is to provide a vehicle for communication between students and staff.

PROGRESS: Interviews with junior college students, faculty members, and administrators, followed by summary conferences, provided the content and guides for the format of the instrument. About 6,500 students in 27 junior colleges across the country responded to an experimental version of the questionnaire. A report (RDR-72-73 No. 3) and a user's manual have been written. Analyses differentiating among colleges of different sizes and among veterans and nonveterans, older and younger students, men and women, and transfer and terminal students have been written.

Chicano Validity Study (CEEB, 142-13) — J. Warren

PURPOSE: Mexican-Americans frequently contend that their bilingual and bicultural background prevents standardized tests from being as valid for them as for members of the majority culture. This contention is being examined by comparing regression equations predicting college performance from test scores and high school grades for Chicano and non-Chicano students in three California state colleges.

PROGRESS: Test scores and high school and college grades have been recorded from the records of the three participating colleges. Grades have been recorded separately by year in college and academic area. Analyses are in progress.

SAT Speededness Study (CEEB, 142-15) F. Evans

PURPOSE: The purpose of the study is to determine the relationship of the aptitude scores obtained under speed and power conditions and the racial and/or ethnic identity of the examinee. The reliability and validity, as well as the level of scores obtained under the various conditions, will be studied.

PROGRESS: Special tests have been administered to volunteer high school juniors in several high schools throughout the United States.

Accuracy of Self-Reported Grades Program (CEEB, 142-16) R. Boldt

PURPOSE: The purpose of this study was to examine the accuracy of student report and of grade-point predictions derived from information supplied by students in connection with the Student Descriptive Questionnaire.

PROGRESS: Data for approximately 719 students from nine American secondary schools were collected. The students were all those at the nine schools who took the SDQ at the first administration, and the schools were systematically chosen for size and diversity.

MAJOR FINDINGS: The correlation between a composite using College Board Verbal and Math along with an actual high school average, and one using College Board Verbal and Math along with the reported high school grades was .97. This fact along with more detailed findings about maximum reporting discrepancies were transmitted informally to the College Board in memoranda.

Evaluation of Preschool for Urban Children Project (Minneapolis Board of Education, 268-6) — D. Norton

PURPOSE: The Minneapolis Board of Education, under a Title III grant, is developing a program to examine ways to enhance the gains made by children viewing Sesame Street. The evaluation efforts are to assess the effects of three different viewing conditions, each representing a different level of parental involvement. In one project component children attend the viewing center with their parents under teacher supervision. Another component consists of children viewing Sesame Street daily at the centers but without their parents. The primary means of communication for these parents is through a social worker and aide. The third component has children and parents viewing Sesame Street in the home. Parents attend weekly meetings aimed to develop a "parental teaching style." Children in all three groups receive support services including follow-up materials, field trip activities, and health services.

PROGRESS: The second operational year (1972-73) continues the Preschool for Urban Children in the same experimental framework as the first year. Pre- and posttests have been administered, and extensive data have been collected monitoring various aspects of parent and child attendance and viewing patterns, project management activities, as well as project health, instructional, and follow-up services. An Interim Evaluation Report was completed in January of 1973. A final report will be completed by July 31, 1973.

MAJOR FINDINGS: A battery of five Sesame Street Cognitive Tests, divided into 25 subtests, were the major assessment instruments used for the evaluation of child achievement. Evaluation results reported in the first year final report (1971-72) indicated that the major product goal for child cognitive achievement was met (with the exception of two minor subtests).

Minneapolis Pyramids Sibling Study, Phase III — D. Norton
(Minneapolis Board of Education, 268-8)

PURPOSE: The Pyramids Reading Project is a major Title I ESEA effort in the Minneapolis Public Schools. Its impact should appear in reading scores of pupils but is hypothesized to be cumulative and not fully assessable for four or more years. The methodology of a sibling study promises better validity for project reports than other evaluative options. A longitudinal study of program outcomes is to be initiated.

PROGRESS: The project was initiated in June 1971. Outcomes of the first phases of the study reveal that older siblings did less well when tested at entry to first grade than did the pupils who received benefits from the Pyramids Reading Program in kindergarten during the 1969-70 school year; the differences that were noted paralleled those found in longitudinal city-wide test data. Phase II for 1971-72 is currently under way.

MAJOR FINDINGS: Initial findings were reported at the 1973 meetings of the American Educational Research Association. The report may be obtained through the ERIC system under the authorship of Daniel P. Norton and Richard W. Faunce.

Efficiency and Efficacy of Evaluation Activities of the State Division of Vocational and Technical Education (State of Illinois Advisory Council on Vocational Education, 268-20) — D. Norton

PURPOSE: The study was designed to furnish the State of Illinois Advisory Council on Vocational Education with data and recommendations on the efficiency and efficacy of the evaluation activities of the State of Illinois Division of Vocational and Technical Education.

PROGRESS: The project was initiated in April 1972 and completed in January 1973. Copies of the report may be requested from: State of Illinois Advisory Council on Vocational Education, 222 South College, Springfield, IL 62706.

MAJOR FINDINGS: The evaluation system used by the Illinois Division of Vocational and Technical Education was found to be perhaps the most effective of its type. However, common deficiencies of all site visitation systems were present. Student-focused follow-up systems were in greatest need if the system was to be improved.

CIRCUS (281-2) S. Anderson

PURPOSE: To develop a comprehensive assessment program for use in nursery school and kindergarten to diagnose the educational needs of individual pupils and to evaluate program effectiveness. The development process will include a national survey of the characteristics of children 4 1/2 to 5 1/2 and the preprimary programs they attend, with special attention to interrelationships among pupil and program variables.

PROGRESS: Based on experience in the ETS-OCD Longitudinal Study of Disadvantaged Children, the evaluation of Sesame Street, and other child development research, 15 measures of children's interests, cognitive and perceptual skills, information processing, problem solving, and divergent production have been developed, along with instruments to assess teacher and program characteristics. These are all designed to be usable by and meaningful to teachers and others without special training in psychology or research. After revisions of the measures on the basis of pilot administrations, two surveys of preschool children and programs were undertaken, one in kindergartens and the other in nursery schools. Representative national samples were obtained, using elementary school sending districts as the primary sampling units. The data collected in these two national administrations are now being analyzed.

A Design for an Accountability System for the New York City School System (New York City Board of Education, 373-1) — F. McDonald

PURPOSE: Increasing desire on the part of the public for more information about the overall effectiveness of its school systems has led in recent years to calls for educators to give an accounting of how well the schools are performing. Underlying these requests are concerns about the efficiency of the schools, the educational progress of the students, and whether all students are receiving opportunities for quality education regardless of their home and community background.

PROGRESS: Early in 1971, ETS was asked by the New York City Board of Education to design an accountability program that could be implemented by the city's public schools. In developing the plan, ETS worked closely with the Board of Education and New York City's Committee on Accountability, appointed by the Chancellor to represent every aspect of public opinion in the city. Among the groups represented in the Committee on Accountability is the United Federation of Teachers, whose collective bargaining agreement in 1969 had called for the development of objective criteria of professional accountability.

In the summer of 1972, ETS submitted a completed design that conceptualized an accountability program as a continuous process for obtaining, interpreting, and using information to improve students' educational performance. The design has three major components: assessment of the performance and progress of students in the schools, diagnosis of differences among schools with respect to students' performance, and corrective action to improve the youngsters' performance and progress.

Within each of these components, the ETS staff made recommendations regarding statistical and methodological models, procedures for collecting and analyzing information, methods for making decisions based on that information, organizational arrangements to achieve the goals of the system, and procedures for monitoring and reporting all phases of the process to the public. Since an effective accountability program must consider a wide array of community, home, and school factors--and must be carried out over a period of time--ETS also developed two time schedules for implementing the program, a short-range, two-year plan, and a long-range, five-year plan.

Advanced Moderator Variable Study (Law School Admission Council, 480-15) — F. Evans, D. Rock

PURPOSE: The purpose of this study was to identify and validate moderator variables which could be used in the prediction of success as a law student.

PROGRESS: A battery of special performance tests (e.g., reading speed, test-taking strategy), a biographical questionnaire, and a personality inventory were administered to approximately 1,000 first-year law students in eight law schools in the Fall of 1969. Using a computer program developed by Rock, groups were formed on the basis of potential moderator variables, and multiple correlations were computed within these groups using standard predictors (LSAT, UGPA) to predict first-year grades. These predictors were then cross-validated and the resultant validities compared.

MAJOR FINDINGS: Students who were somewhat older than the average first-year student were found to be more predictable; i.e., a higher validity was observed. Also, students from higher socioeconomic backgrounds tended to be relatively unpredictable as well as somewhat underpredicted. Further analysis designed to identify "nonstandard" predictors (other than LSAT, Writing test, and UGPA) for the less predictable groups met with limited success. The study is reported in the Law School Admission Council 1973 Annual Report.

Factor Analysis of Law School Grades (Law School Admission Council, 480-47) R. Boldt

PURPOSE: Proliferation of novel courses in the latter stages of law school education has led to a concern that new abilities are needed for successful performance in law school. While it is reasonable to suppose, though certainly not empirically justified, that law school performance in the past was based primarily on verbal skills, one wonders whether purely verbal skills are currently enough. Hence, an investigation of the factor structure was undertaken. Because of the opportunity afforded, data on instructors were also collected.

PROGRESS: Grades of 116 law school students were obtained from transcripts, along with LSAT, Writing Ability, and Background scores. Analyses of one- through four-factor systems indicated that the system of grades was based on essentially one factor. The residuals resulting from the factor analysis were used to study the possible existence of instructor bias. Two instructors were found with residuals significantly different from zero; one was found to be a slightly easy grader, the other somewhat harder. These results are reported in RB-73-42.

Validity Study of Experimental Item Types
(Law School Admission Council, 480-51)

B. Pitcher
A. Carlson
M. McPeek

PURPOSE: The validity of several experimental item types for predicting first-year average grades in law school is being investigated in comparison with the validity of the item types currently being used in the operational sections of the test.

PROGRESS: A tape combining test score and validity study records into a single record for each of more than 15,000 students from 90 law schools has been prepared for use in this study. Scale conversions of test scores and grades have been completed so that data may be pooled across law schools.

An Exploratory Study of the Possibility of Curvilinearity in Law School Validity Study Data (Law School Admission Council, 480-57)

B. Pitcher

PURPOSE: This study is a preliminary, exploratory attempt to put together some data to see whether it would be useful to pursue further the question of curvilinearity in law school validity study data--more specifically, to see whether the relation between Law School Admission Test scores and first-year average grades in law school (FYA), the relation between undergraduate grades (UGPA) and FYA, and the relation between a Predictive Combination based on LSAT scores, UGPA, and Writing Ability scores and FYA, appear to depart from linearity enough to make the question worthy of further investigation.

PROGRESS: Data from 15 law schools that participated in the 1968-69 validity studies were used to prepare plots for an exploratory inspection of the data. A report of this preliminary work is being prepared for distribution to the LSAC, and a proposal to do further investigation of the question of curvilinearity has been requested.

MAJOR FINDINGS: The data for 30 law school groups--two classes at each of 15 schools--were used to test one specific hypothesis, namely, that students whose test scores and undergraduate grades place them at the low end of the scale tend to do better than predicted, and conversely, students at the upper end of the scale earn grades lower than predicted. This hypothesis was supported by the data for the Predictive Combination based on UGPA, LSAT, and Writing Ability, at both ends of the scale, but not for LSAT nor for UGPA when considered separately.

Validity Study of LSAT Part-Score Reporting — B. Pitcher
(Law School Admission Council, 480-59)

PURPOSE: This study will investigate the validity for predicting first-year average grades in law school of several part-scores derived from the separate section scores of the Law School Admission Test. All analyses are to be done separately by law school. Included within this study is a comparison of the validity of the traditional predictors, undergraduate average, LSAT and Writing Ability scores, for male and female law students at schools that have at least 50 female students in their 1971 entering classes.

PROGRESS: A sample of 15 law schools has been selected from the group of 99 law schools that participated in the 1972-73 Law School Validity Study Service, and a tape containing the data for the sample has been prepared for use in this study.

Student Subgroups Validity Study — B. Pitcher
(Law School Admission Council, 480-60)

PURPOSE: This study is designed to compare the prediction of first-year average grades in law school from the traditional predictors, undergraduate average grades, Law School Admission Test and Writing Ability scores, for various subgroups of students pooled across law schools.

PROGRESS: Data for 90 law schools have been extracted from the tapes used for the Law School Validity Study Service in 1972-73, first-year average grades have been converted to a common scale, and a tape for use in this study has been prepared. Preliminary counts indicate that there should be enough data for comparing groups based on sex, ethnic origin, age, and undergraduate major, although some of the planned subcategories may not be usable. The statistical analysis of the special subgroup of students whose records included more than one set of LSAT and WA scores is essentially completed.

Effect of Differences in College Grading Standards on the Prediction of Law School Grades (Law School Admission Council, 480-61)

W. Schrader
B. Pitcher

PURPOSE: To find out whether the mean performance on various measures for law students who attended a particular college would be useful supplementary predictors of law school grades. In this study, college means on LSAT scores, on Writing Ability scores, on undergraduate average grades, and on law school grades for the college attended by a student were treated as possible predictors of the student's law school performance.

PROGRESS: College means, each based on 10 or more law students, were calculated for 125 colleges, using data for 38 law schools. The resulting means were used along with Undergraduate Grade Point Average, LSAT scores, and Writing Ability scores to predict first-year law school grades. In 13 law schools, shrunken multiple correlation coefficients involving one or more of the college means along with the three usual predictors were compared with the shrunken multiple correlation coefficient based on the three predictors. In 7 law schools, data for two successive entering classes were available. In these schools, regression weights for students entering in one year were used for predicting grades of students who entered the law school in the preceding or following year. Validity coefficients for combinations involving one or more college means as well as the three usual predictors were compared with validity coefficients based on the three predictors. A report of the study is being prepared.

MAJOR FINDINGS: Although results varied from one law school to another, as would be expected, none of the college means or combinations of college means contributed substantially to prediction when median validity coefficients for each group of law schools were considered.

Teacher Behavior Research (523-6) F. McDonald

PURPOSE: To explore the relationship between teacher behavior and student outcomes. This is a program of research directed toward the development of performance-based teacher training and assessment procedures.

PROGRESS: Teacher sessions have been designed in which the length of lesson, content taught, type of student taught, and teaching strategies are identical. Two types of sessions have been developed--a shorter one, only 20 minutes in length, taught to six students, and a longer one, in which the teacher teaches a lesson of regular length to a class of typical size every day for one or two weeks.

Audio and video recordings are made of the teachers' behavior; pupils rate the teachers and, in the longer courses, take an achievement test. The teachers also rate their own teaching performance.

The research goals are to establish the reliability and validity of these procedures, and to study the relation of teaching performance to student learning and the training variables which produce the acquisition of teaching skills. In addition, a related line of research which is being conducted in conjunction with the Personality and Social Behavior Research Group under H. A. Witkin is examining the relationship between cognitive style, teacher behavior, and student behavior and outcomes.

Finally, a film test or simulation device designed to measure teachers' perceptions of the classroom situation, their sensitivity to student behavior, as well as other dimensions, has been developed and pilot tested with more than 350 teachers. A predictive validity study is now under way at Temple University; initial results support the promise of the instrument both as an assessment technique and as a training device.

National Commission on Performance-Based Education (Rockefeller Brothers Fund, 525-1) — F. McDonald

PURPOSE: To assess the feasibility of creating a national organization to coordinate and lead the movement to performance-based teacher education and certification.

PROGRESS: One of the major purposes of this study was to determine what a national organization ought to do. In the course of this study we spoke to practically everyone now doing something in performance-based teacher education--program directors, designers of instructional systems, and state department officers planning changes in the requirements for certification. Our purpose in these discussions was to find out what was being done and what they thought needed to be done.

The most fundamental need to emerge was to establish what competencies a teacher ought to acquire to produce significant effects on student learning. The lack of valid descriptions of the competencies to be acquired and the diversity of those which have been chosen as the goals of developing teacher education programs have made assessment and evaluation almost impossible. Assessment and evaluation procedures are badly needed, but such techniques cannot be developed without knowing more precisely what is to be measured. Moreover, programs for assessing the competencies of teachers from different institutions cannot be created as long as there are diverse concepts of what constitutes competence.

Other problems which emerged, as we talked with those individuals most responsible for managing new programs, involved the management problems which arise from the radical differences occurring between traditional teacher education and performance-based programs. In addition, instructional modules or systems for training are tied to a particular developer's concept of teaching competence and the work on these systems is in a relatively primitive stage. Few systems have undergone adequate tests of their effectiveness.

The picture that emerged from this study was one in which many individuals were making great efforts to build performance-based programs. But the tasks to be mastered were beyond the capabilities of individuals or even of a single institution. Moreover, the movement appears fragmented with poor internal communications. A national organization is needed to mobilize and coordinate present efforts, to start meaningful programs of research and development on what constitutes competence, how to measure and train for it, and to provide practical experience in and training for managing programs.

MAJOR FINDINGS: This study resulted in a proposal for the foundation of a National Commission on Performance-Based Education. The major purpose of this Commission will be to create a research and development program directed to four goals: (1) to identify through systematic research those teaching skills which if acquired and used constitute teaching competence; (2) to develop measures of these skills and assessment strategies for evaluating competence; (3) to develop and test instructional systems and their components which facilitate learning these skills and achieving competence; (4) to develop management systems and provide training for the directors and administrators of performance-based programs.

GRE Bayesian Validity Study (GRE Board, 540-64) R. Boldt

PURPOSE: The purpose of this study is primarily to gain experience in the application of a Bayesian model in the validity study setting and secondarily to accomplish a graduate school validity study using GRE variables. The model used allows the pooling of data in a situation where criterion observations may be scattered over a variety of institutions where they are not necessarily on the same scale. The model is being proved against a least squares model which allows a partial pooling of data.

PROGRESS: Data were solicited from departments of psychology and economics that require the GRE and the Advanced Tests for admissions but insufficient data were received. Data from the College Board Validity Study Service have been used instead.

MAJOR FINDINGS: The least squares and Bayesian systems are almost exactly equivalent in cross-samples from a correlational point of view. However, average errors of estimate indicate a slight superiority of the Bayesian system. It is speculated that this superiority arises from a more comprehensive pooling afforded by the Bayesian system.

Graduate Performance by Critical Incident (GRE Board, 540-75) R. Reilly

PURPOSE: The purpose of this project is to better define the criteria by which graduate students are evaluated through the use of a list of Behavioral Descriptions supplied by graduate faculty. The secondary objective is to contrast the relative importance of different criteria across three discipline areas representing the broad spectrum of graduate disciplines. Faculty from the areas of Psychology, Chemistry, and English utilized the list of the Behavioral Descriptions to describe the performance of graduate students with whom they have had contact.

PROGRESS: Data collected from 1300 faculty members have been analyzed and a report prepared for the Graduate Record Examinations Board.

MAJOR FINDINGS: Factor analyses revealed a total of eight factors which were relatively consistent across the three disciplines. Behaviorally anchored rating scales are currently being developed based on the factors.

Word Associations of Black and White College Students (GRE Board, 540-77) — J. Campbell

PURPOSE: This project was designed to compare word association responses for college students from four different kinds of colleges in three different geographic areas: (1) selective predominantly white colleges, (2) less selective predominantly white colleges, (3) selective predominantly black colleges, and (4) less selective predominantly black colleges. Stimulus words were chosen from those used in verbal analogy and antonym items.

PROGRESS: Responses on 200 words have been obtained from approximately 3,000 students. Protocols have been edited and keypunched preparatory to data analysis. (Study being done in collaboration with Dr. Leon Belcher of Texas Southern University.)

Women Doctoral Recipients (GRE Board, 540-78) J. Centra

PURPOSE: The main purpose of this study is to describe the current status and professional development of a sample of women who have received a doctorate during the past 20 years or so. Trends in their experiences and backgrounds will be compared with a sample of male doctorates matched by a field of study and graduating institution. A secondary purpose of the study is to obtain information on the views toward their own graduate training of both females and males.

PROGRESS: Over 5000 graduates from the years 1950, 1960, and 1968 were selected. Among the areas covered in the study were the so-called "professional socialization" process, patterns of interruptions and obstacles to career development, demographic characteristics, and reentry problems. The questionnaire was mailed out during the Spring of 1973, and a final report is expected by December 1973.

Effectiveness of the GRE Fee-Waiver Program (GRE Board, 540-79) — R. Feldmesser

PURPOSE: The GRE board has established a program under which low-income seniors are permitted to take the GRE without payment of the test fee. This study sought to learn how effective the program was in reaching the students for whom it is intended, encouraging them to consider attendance at graduate school in planning their post-baccalaureate careers, and affecting the composition of applicants and those admitted to graduate school.

PROGRESS: Questionnaires were sent to most 1971-72 fee-waiver recipients and were returned by a large proportion. A number of local program coordinators were interviewed. Data have been analyzed, and a final report has been submitted to the GRE Board.

MAJOR FINDINGS: Students who received the waivers did come heavily from low-income and minority-group families. However, communications obstacles, and the relatively small amount of money directly involved, prevented the program from being used to the fullest extent possible.

GRE-Quantitative Coaching Study (GRE Board, 540-83) F. Evans

PURPOSE: A major purpose of this study is to design and test a special program of instruction to increase scores on the Graduate Record Examinations Quantitative Aptitude Test. The effectiveness of the special instructional program will be evaluated for groups of Black, Chicano, and Caucasian college seniors. The study is divided into three parts: feasibility (to determine the need and interest in a special instructional program); exploratory (to identify specific areas to be covered by instruction); and operational (to design and produce a special curriculum and evaluate its effectiveness).

PROGRESS: The feasibility study and exploratory phase have been completed. Materials have been developed, and special instruction for GRE-Q will be offered on several campuses in the Fall of 1973.

The Role of Cognitive Style in Students' Academic Evolution from High School Through Graduate and Professional School (GRE Board, 540-86; NIMH, 737-2)

H. Witkin
P. Oltman
D. Goodenough

PURPOSE: One aim of this study is to assess the role of cognitive styles in students' academic choices and performance in the college period, selection of elective courses and majors, shifts in major, performance in various subject matter areas, post-college aspirations. A second aim is to assess the role of students' cognitive styles in predicting the decision to go on to graduate or professional school; the kind of school favored, eligibility for graduate or professional school, specialty selected after admission, and performance in graduate or professional school. A third aim is to explore continuity in students' choices and performance, as a function of cognitive style, from the high school through the higher-education levels.

PROGRESS: A four-year longitudinal study, through the college years, of an entire class of a large municipal college, has already been completed. Cognitive style data were obtained at admission for all students in the class and a more intensive study was made of smaller, selected groups showing different kinds of cognitive patterns. With the help of the GRE Board we have identified the graduate schools now being attended by those students in the original college sample who made the decision to go on to graduate school. Similarly, through the cooperation of the Law School Admission Council, the American Association of Dental Schools, and the Association of American Medical Colleges, we are in the process of identifying the professional schools now being attended by those of our students who made the decision to go on into these professions. Finally, college and high school transcripts have already been obtained for the entire college class.

Some of the results of the study from analyses already completed have been summarized in Research Bulletin 73-11.

Development of Provisional Criteria for the Study of Scientific Creativity (GRE Board, 540-87)

N. Frederiksen
W. Ward

PURPOSE: The aim is to develop tests of creative performance that may be characterized as criterion measures and that subsequently may be useful in studies of the creative performance of scientists. The tests will be based on job samples or simulations of the work of a behavioral scientist and will attempt to measure abilities required of a scientist in conceiving, planning, and evaluating research studies and in interpreting research data. Reliability and validity will be investigated. If the tests prove to be satisfactory psychometrically, they should be useful in investigating the correlates of various aspects of creative performance and providing dependent variables in experiments on situational factors that might influence the behavior of a scientist.

PROGRESS: Tests have been or are being developed with the following titles: Formulating Hypotheses, Evaluating Hypotheses, Evaluating Proposals, Measuring Constructs, Analyzing Constructs, Solving Methodological Problems, Formulating Research Ideas, Evaluating Manuscripts, Personnel Selection Problems, and Ideational Fluency in Psychology. The Formulating Hypotheses test, which contains 18 new items based on behavioral science studies, has been pretested with graduate students at the University of Pennsylvania. Data have been collected at the University of Washington in an experiment designed to solve a methodological problem pertaining to the influence of certain instructions on quality of performance. Plans have been made for the tryout of four tests in a nationwide administration of a test to several thousand students, using an item-sampling approach.

National Assessment (553) J. Fremer

PURPOSE: This is a project of the Education Commission of the States (ECS) directed by the staff of National Assessment of Educational Progress (NAEP). The purpose is to assess the attainment of educational goals in selected subject-matter areas by a representative sample of 9-, 13-, 17-year-olds and young adults, and to measure progress over time by periodic reassessments.

PROGRESS: ETS has been contributing to the National Assessment project since its preliminary planning phases in 1964-65. Our early work included explorations of the goals in nine subject-matter areas and a number of special studies. Feasibility studies were undertaken of a variety of exercise (test item) formats; these included some designed to demonstrate a skill and many which required unusual stimuli and response modes. Other studies centered around the development of "easy tasks" to determine if it were possible to construct valid measures of what 90% or more of a group has achieved. Still other studies looked for ways to avoid cultural, sex, SES, geographic, and other biases. ETS has developed exercises in six assessment areas: writing, literature, science, social studies, music, and art. (Other organizations have developed materials for reading, mathematics, career and occupational development, and citizenship.)

The schedule for administering National Assessment Exercises is as follows:

Assessment Year		Subject Areas
01	1969-70	Writing, Citizenship, Science
02	1970-71	Reading, Literature
03	1971-72	Music, Social Studies
04	1972-73	Mathematics, Science
05	1973-74	Career and Occupational Development, Writing
06	1974-75	Art, Citizenship
07	1975-76	Reading, Literature
08	1976-77	Repeats year 03, future years repeat areas covered five years previously

During the period July 1, 1972 through June 30, 1973 ETS completed work on exercises for the second cycle of the writing assessment, the year 05 (1973-74) administration. In addition, exercise development was initialed for the second cycle of the Literature assessment, scheduled for year 07 (1975-76) administration. Work began on January 1, 1973 and, following planning conferences and exercise writer workshops, literature exercises for ages adult, 17-year-old and 13-year-old were developed and field tested by June 30, 1973. Post field trial reviews by subject and lay committees were also carried out for the adult and 17-year-old exercises.

During the July 1, 1973 to June 30, 1974 year, age 9 exercises will be developed, field tested, and reviewed. Exercises for all age levels will receive further revisions and trials as is called for by internal and external reviews. Further exercise development to complete coverage of objectives is also anticipated.

Mr. Michael Ward, Miss Gena Reisner, and Mrs. Linda Conrad have major responsibility for the National Assessment of Literature with the assistance of Mr. Robert Atwan, Miss Trudy Conlan, Mrs. Ann Groom, Mr. Robert Jones, and Mrs. Geraldine Smith.

MAJOR FINDINGS: NAEP has written numerous brochures and articles as well as reports of the results of the assessments that are already completed. A list of publications can be obtained from the office of the National Assessment Information Service.

Comparison between Faculty Peer Ratings and Student Ratings (572-72) — J. Centra

PURPOSE: To investigate the relationship between peer ratings of teaching based on classroom visits and the ratings given by students.

PROGRESS: Peer ratings, based on two visits by each of three faculty members, are being compared to the mean ratings of students in each class on 16 items dealing with teaching effectiveness. The final report should be available by February 1974.

The Relationship between Student and Alumni Ratings of Teachers (572-72) — J. Centra

PURPOSE: To investigate further the relationship between student and alumni ratings of instructors.

PROGRESS: A report has been issued (RB-73-39).

MAJOR FINDINGS: Student and alumni ratings for 23 teachers were found to correlate .75 (somewhat less for teachers rated only by graduates of their department). This substantial agreement between current students and alumni (of five years) regarding who have been effective or ineffective teachers suggests that student ratings are fairly permanent and do, at least in part, reflect overall, long-term effects of instruction.

The Physical Activities Survey of Police Officers in New Jersey (574-1) — L. Goldstein

PURPOSE: To collect information about the kinds of physical activities police officers perform, their present health status, what measures they take to maintain good physical condition, and their appraisal of the present civil service physical performance test battery in order to assist the Department of Civil Service in an evaluation of the efficacy of the constituent parts of the present physical performance test, to uncover any significant relations between job requirements and the maintenance of good health, and to signal the need for specific action that the Department, State Law Enforcement Planning Agency, or other state or local agencies might take in order to assist police officers in maintaining good physical condition. (Project sponsored by the State of New Jersey, Department of Civil Service and State Law Enforcement Planning Agency.)

PROGRESS: A Physical Activities Survey questionnaire was developed, field tested, and distributed to the police departments of 132 communities in New Jersey which use civil service procedures for the selection of their police officers. Questionnaires completed by men at all ranks from patrolman through chief in 92 jurisdictions comprise the data for this study. A report has been prepared (PR-73-20).

MAJOR FINDINGS: The distributions by county and rank of the 1875 respondents are representative of the larger distributions of police officers in the state. The typical New Jersey police officer, as defined by the PAS sample, is approximately 37.5 years old, stands just under 71 inches, weighs almost 189 pounds and has been on the job for 11 1/2 years. He has gained more than 14 pounds since appointment but describes his present physical condition as "good." Nevertheless, 57% of the group claims at least one physical symptom or disease while more than 28% have two or more complaints. Back trouble, loss of teeth, and hemorrhoids are those conditions most frequently mentioned (16-24% of the group); high blood pressure, ulcer and nervous disorder (7-9%) are the next most common complaints.

More than three-fourths of the group identify sitting as the physical activity required most on the job. Walking is an activity performed by more than half of the group while running is listed by less than 12%. There is some evidence which indicates that diminishing health status for this sample is related not only to age but also to the sedentary nature of the job. Extreme weight gain appears to be implicated in increased complaints of diseases or symptoms, too.

Among others, a suggestion is made that programs of regular exercise and of weight control, instituted at the local jurisdiction level with the aid of federal or state funding, might be effective for the maintenance or improvement of police officers' physical condition.

Promotion Examinations for Police Sergeants, Lieutenants and Captains in New Jersey (574-1) — L. Goldstein

PURPOSE: To assist the Department of Civil Service in the development of improved promotion examinations for sergeants, lieutenants, and captains in New Jersey police jurisdictions by collecting and analyzing data relating the kinds of duties these superior officers perform to the size of the community in which the jurisdiction is located. (Project sponsored by the State of New Jersey, Department of Civil Service and State Law Enforcement Planning Agency.)

PROGRESS: A booklet of behaviorally stated objectives has been developed for each rank. Approximately 35-40 performances are to be rated for frequency and importance by men below, at, and above the rank assigned to police departments in communities ranging in size from 100,000 and over to less than 10,000. ANOVA techniques will be used in order to determine significance of main effects of rank of rater and community size and of the interaction. Those behaviors for men at each of the three specified ranks which are common to all jurisdictions irrespective of size will be identified in terms of their frequency and importance. Those behaviors which are performed more frequently or are more important in some communities as compared to others will also be differentiated. In this way, a group of behaviors will be isolated, for each rank, to serve as a basis for items for future promotion examinations.

Personality Structure and Development in Headstart Children (OCD, 581-14; 779-1) — W. Emmerich

PURPOSE: As part of the Longitudinal Study of Disadvantaged Children and Their First School Experiences, this study has the following aims: (1) Evaluation of a multivariate observation procedure for deriving personal-social measures on young children during relatively unstructured class periods; (2) Discovery of the structure (organization) of personal-social behaviors in Headstart children and possible variations in structure associated with the child's sex, age at entry, and first versus second semester enrollment; (3) Examination and interpretation of mean differences on personal-social constructs associated with the child's sex, age at entry, first versus second semester enrollment, and their interactions; (4) Using a personality typology based upon the structural analyses, stability and change in personal-social behaviors during the Headstart year will be evaluated. A model for predicting the qualitative directions of such change will be tested; (5) The classroom personal-social measures will be related to independently assessed measures of cognition, style, and personality in the child; (6) It is hypothesized that parent-child and teacher-child relationships will jointly influence the child's personal-social behaviors in the classroom. These hypotheses will be tested by examining main effects and interactions among maternal and teacher variables with regard to their impact upon the classroom personal-social child measures; (7) The personality types will be conceptualized and tested empirically as moderators of functional relationships between educational experiences (e.g., teacher classroom behaviors) and cognitive development (e.g., test performances at the end of Headstart).

PROGRESS: All Headstart period data have been collected, and those bearing on aims 1-4 have been coded and are being analyzed. A report on aims 1-4 is found in "Disadvantaged Children and Their First School Experiences: Structure and Development of Personal-Social Behaviors in Preschool Settings" (PR-71-20).

Preschool Teachers' Beliefs on Effective Teaching Techniques and Their Relationships to Pupil Characteristics (OCD, 581-14)

W. Emmerich

PURPOSE: The Enhancement of Learning Inventory (ELI) was constructed to assess a teacher's beliefs about the effectiveness of a variety of methods for teaching each pupil in her classroom. The present study examined these beliefs in relation to pupil cognitive and background characteristics for 35 teachers of economically disadvantaged preschool children, most of whom were enrolled in Head Start.

MAJOR FINDINGS: Individual differences among pupils prior to preschool entry were related to teachers' beliefs about effectiveness of teaching techniques toward the end of the preschool year. For example, pupils whose mothers had more years of school were judged by teachers as generally more receptive to classroom learning. However, processes engaged by classroom learning (cognitive skills and style) were more strongly and consistently related to teacher beliefs than were the child's family-background characteristics. Pupils exhibiting relatively high levels of cognitive skill prior to preschool were judged by teachers generally as most capable of classroom learning, especially when pupils were given considerable autonomy in structuring their classroom learning experiences. Also, children with moderate response tempos were perceived as more receptive to teaching efforts than children with either fast or slow response tempos. The child's sex and age at preschool entry also influenced teacher judgments. Different patterns of teaching style and teacher individuation were related to the pupil's initial characteristics and in several instances also to cognitive growth during the preschool year. While there was evidence that teacher belief patterns influence certain aspects of pupil psycho-educational development, the findings generally indicate that pupil behavioral characteristics may have a greater impact upon teacher behavior than vice versa.

A report is currently in preparation (PR-73-12).

ETS-Head Start Longitudinal Study of Disadvantaged Children and Their First School Experiences
(OCD, 581-14, 581-23)

V. Shipman

PURPOSE: The purposes of the Longitudinal Study are to identify the components of early education that are associated with the cognitive, personal, and social development of young children; to determine the environmental and background modifiers which influence these associations; and to describe and evaluate the interaction of these as part of the young child's development over a period of seven years (from approximately age 3 1/2 through third grade). The study has been sponsored and supported by the Department of Health, Education and Welfare, Office of Child Development (Grant #H-8256).

PROGRESS: Major activities of the past year (Year 5 of the study) have emphasized reporting data collected in previous years. To this end, several reports were completed, covering a variety of concerns including the extent of structural stability and change in child test scores between Years 1 and 2 (PR-72-18), a summary report of the findings therein (PR-73-8), and a series of technical reports for measures used during the first two years of the study (PR-72-27). In addition, separate reports were completed for (1) preschool teachers' beliefs on effective teaching techniques and their relationship to pupil characteristics (PR-73-12); (2) the development of self-regulatory behaviors (PR-73-18) and the characteristics of the urban preschool centers sampled as determined by the Preschool Center Inventory (PR-73-21). Further, findings from the Year 1 Eight-Block Mother/Child Interaction task will also be reported in the coming weeks. Several additional reports are scheduled for completion over the next few months.

Coding, processing and analyzing of collected data has also continued. Specifically, almost all child test data through Year 4 (1971-72) have been processed.

Data collection during the past year has been minimal to allow for a concerted reporting effort. However, group achievement tests were administered and classroom observations and teacher ratings were obtained. The last scheduled data collection will be conducted in FY '74 when most study children are to be in third grade; current plans call for extensive individual and group child testing, home interviewing, and the collection of various teacher ratings, classroom observations, and school and community information.

MAJOR FINDINGS: See PR-71-19, "Disadvantaged Children and Their First School Experiences: Structure and Development of Cognitive Competencies and Styles Prior to School Entry," December 1971; PR-71-20, "Disadvantaged Children and Their First School Experiences--Structure and Development of Personal-Social Behaviors in Preschool Settings," November 1971; and PR-72-13, "Demographic Indexes of Socioeconomic Status and Maternal Behaviors and Attitudes," June 1972. Additionally, please note reports listed above.

Audit of Project IMPACT (Phase II) (Title I, 585-41) — D. Powers

PURPOSE: To provide an independent check on the activities conducted for the evaluation of Project IMPACT--Duval County, Florida's performance-contracted teacher support program.

PROGRESS: All audit activities have been completed and a final report (PR-73-10) has been submitted. Audit activities included a review of each of the following: the goals of the project, the project's evaluation design, test administration procedures, test content, data collection procedures, data analysis procedures, and evaluation reports.

MAJOR FINDINGS: Aspects of the evaluation in which weaknesses were perceived were reported. Also noted were those areas in which improvements had been made from the previous year.

Evaluation of the Virginia Beach Extended School Year Program (585-60) — D. Powers

PURPOSE: To determine the effects of a year-round education plan on the achievement and attitudes of the children involved in a three-year pilot program.

PROGRESS: Baseline achievement and attitude data have been gathered for a sample of students enrolled in four pilot schools and two control schools in the Virginia Beach Public School System. Additional data will be collected on the same students after the year-round plan has been in effect for one year and two years, respectively. Teachers' ratings of the degree to which their objectives correspond to the items contained in the achievement battery have been obtained and will be used to differentially weight sets of items. Cross-sectional data obtained from the state testing program will also be used to assess program effects. A progress report which describes first-year activities and presents baseline data is being prepared.

Evaluation of the New Approach Method (NAM) D. Powers
(OCD, 589-10)

PURPOSE: To determine the effectiveness of an experimental program, the New Approach Method (NAM), in teaching preschool-aged children beginning reading. The project has been supported by funds from the U. S. Office of Child Development.

PROGRESS: A battery of instruments was assembled or adapted and administered by a community-based field staff to children enrolled in the NAM program at five Trenton, New Jersey day-care centers, as well as to a randomly selected control group at one of the centers. Questionnaires were developed and administered to the parents of the children involved in the program. Additional data regarding program operations were also gathered.

Data collection activities have been completed, and the data from these efforts are currently being analyzed. A preliminary report has been submitted, and a final report will soon be available.

Survey of Adult Reading Performance (NIE, 592-2) R. Murphy

PURPOSE: A set of real-life reading tasks is being used to assess the reading skills of American adults. The tasks were developed on the basis of information obtained in a previous national survey of adult reading activities. A project report (PR-72-24) by Amiel Sharon describes the results of the activities survey. The survey of adult reading performance will provide difficulty levels on approximately 170 items for a national probability sample and for a number of selected subgroups.

PROGRESS: The reading tasks are being administered nationwide by Response Analysis Corporation to a sample of 7500 adults 16 years old or older. In addition to responses to the reading tasks, information such as age, income, and educational background is being collected on the sample.

MAJOR FINDINGS: The reading activities survey showed that American adults spend a considerable amount of their time in reading while working, traveling, and shopping and that they consider such reading activities important. The national performance survey will assess the extent to which various groups of adults can read materials intended to faithfully reflect these important reading activities.

A Descriptive and Analytic Study of Compensatory Reading Programs (USOE, 593-1) — D. Trismen

PURPOSE: In July 1971 ETS was requested by the U. S. Office of Education to develop design and analysis plans for a study of compensatory reading programs in grades 2, 4, and 6 of the U. S. public schools. The overall purpose of the study is to describe and to document the effectiveness of the various reading programs that are offered under the heading of compensatory education.

PROGRESS: The planning activity for the study took place during the period July-December 1971 and is described in the ETS Final Report for Contract No. OEC-0-71-3715. (A Descriptive and Analytic Study of Compensatory Reading Programs, January 1972.) The first phase of the actual study was carried out during the school year 1971-72. (A report on that phase will be available by July 1.) Phase I consisted of a questionnaire survey in Spring 1972 of reading programs (compensatory and noncompensatory) offered in grades 2, 4, and 6 of a national sample of over 700 public schools. Data were gathered descriptive of schools, teachers, students, and instructional practices. From the Phase I data were then chosen 250 schools that were invited to participate in Phase II: 100 schools with compensatory reading programs funded by Title I; 100 schools with compensatory reading programs funded by sources other than Title I; and 50 schools offering no compensatory reading programs at all. An additional 30 schools with programs judged to be particularly innovative were also included. Data collection consisted of administering the same questionnaires that were administered in Phase I; pretesting and posttesting all students in grades 2, 4, and 6 using a reading achievement battery and a measure of attitudes toward reading; keeping records of students' attendance at reading instruction; and, finally, in the innovative schools, conducting a series of classroom observations. The analysis of Phase II data will involve relating achievement and attitude gain scores to program type in an effort to assess the relative effectiveness of the various programs.

A summer program study is under way and will continue through the summer of 1973. Phase III, planned for the school year of 1973-74, will include a closer look at about 50 of the schools included in Phases I and II. For each program type delineated (see Major Findings below), schools judged to be most and least effective will be visited. Staff will be interviewed and classroom observations conducted in an effort to document in detail the characteristics of reading programs that have been found to be most and least effective in terms of the criteria of this study.

MAJOR FINDINGS: Analysis of the Phase I (descriptive) data has just been completed. Five major program types were identified from the data provided by principals and teachers in questionnaires. Detailed descriptions of the program types, the schools they occur in, and the students they are intended for will be found in the final report for Phase I.

Survey of Disparities in Educational Achievements — R. Feldmesser
(Ford Foundation, 595-1)

PURPOSE: This is an effort to determine whether existing data indicate any trends in the disparities in educational achievements between white and minority-group students since the time the Equal Educational Opportunity Survey (EEOS) was conducted in 1965.

PROGRESS: The research literature and reports of school-district and statewide testing programs have been searched for measures of disparities comparable to those in the EEOS, and both these measures and the EEOS data have been recalculated where necessary to maximize comparability. The final report is in preparation.

MAJOR FINDINGS: Insofar as comparable data are available, they suggest that disparities have increased since 1965, though the trend may have been reversed in the last two years. However, the data are not sufficient in either quantity or quality to permit trend statements to be made with confidence.

Learning System in the Social Sciences (Title III, 597-10) — R. Murphy

PURPOSE: Mercer County Community College is designing and developing an experimental program in the social sciences as the initial program for its learning resources center. The purpose of the project is to develop an experimental program that will employ instructional techniques that are effective for minority students, involve a curriculum that is relevant and purposeful to minority students, and provide opportunities for strengthening the basic academic skills.

PROGRESS: To date, draft materials have been produced and have been reviewed by students and adult members of the Trenton community. Implementation of the program will begin in September 1973.

A Survey of Skills and Abilities Needed for Graduate Study in Business (Graduate Business Admissions Council, 693-59)

J. Campbell
P. Casserly

PURPOSE: This project was designed to investigate the cognitive abilities necessary for student success in graduate schools of business.

PROGRESS: Job analysis type interviews were held with students, faculty members and administrators at 19 graduate schools of business. In addition to activities drawing on cognitive abilities, the interviews discussed affective propensities of students, their undergraduate preparation and work experience, and a number of other factors.

MAJOR FINDINGS: (1) The verbal and quantitative areas presently measured by the ATGSB are essential, (2) Lack of knowledge in mathematics causes difficulty for students in many schools, (3) Ability to write clearly and simply is an essential skill, (4) Some students who are low on "tolerance for ambiguity" have difficulty with business problems where some important factors cannot be reduced to known, quantifiable variables, (5) Other cognitive factors including general reasoning, inductive reasoning, and deductive reasoning are undoubtedly important, but appear to be adequately represented in the present composition of the ATGSB, (6) Motivation is of predominant importance. Several possibilities for assisting the schools to obtain better assessment of students' motivation should be explored. A draft report has been submitted to the Graduate Business Admissions Council Research Committee.

The Effect of Stress on Dream Affect (NIMH, 737-2)

D. Goodenough
H. Witkin

PURPOSE: This study had two major goals: (1) to study systematically the consequences of viewing stress films on the emotional content of dreams and (2) to test the hypothesis that respiratory correlates of emotion are similar in waking and dreaming states.

PROGRESS: Data have been collected on 28 subjects who slept in the laboratory after viewing stress and neutral films and were awakened from REM sleep to collect dream reports. A manuscript describing the study has been completed.

MAJOR FINDINGS: Anxiety in dreams reported for subjects was significantly increased on nights following stress film viewing. Emotion-respiration relationships appeared to be congruent across waking and sleeping states.

Genetic Factors in the Development of Field Dependence and Spatial Abilities: Intercorrelations Among Family Members (NIMH, 737-2)

D. Goodenough
H. Witkin

PURPOSE: It has been proposed that certain cognitive abilities have, as one determinant, a gene located on the X chromosome. If true, then one might expect mother-son and father-daughter correlations of about equal magnitude; lower father-son correlations; and mother-daughter correlations somewhere between these two values. Previous studies have confirmed these deductions for tests of visualization, but are contradictory for measures of field dependence. An attempt is being made to replicate and extend this work.

PROGRESS: A battery of cognitive tests has been administered to a number of families with father, mother, son, and daughter. Data analyses have begun.

Genetic Factors in the Development of Field Dependence and Spatial Abilities: X-Linkage Study (NIMH, 737-2)

D. Goodenough
H. Witkin
I. Olkin

PURPOSE: It has been proposed that certain cognitive abilities have, as one determinant, a gene located on the X chromosome. If true, then in patterns of family transmission one might expect these abilities to show linkage with characteristics that are known to be determined by genes on the X chromosome (marker genes). An attempt is being made to test this hypothesis for a number of cognitive factors including field dependence, space, and visualization.

PROGRESS: A battery of cognitive tests has been administered to three sons for each of a number of families. Marker gene data have been collected for two blood types (XGa and G6PD) and for two types of red-green color blindness (protan and deuteran). Analyses of the data have begun.

Respiratory Correlates of Emotional States
(NIMH, 737-2)

D. Goodenough
P. Oltman

PURPOSE: Respiratory rate (or breath time) is not very clearly related to emotional states. It has been proposed that the breath cycle is composed of three relatively independent components (inspiration, expiration, and post-expiratory pause), and that a clearer understanding of the respiratory correlates of emotion might be had if these components are examined separately. Many hand scoring attempts to measure these components have been made in the history of research in this area, but these attempts have been largely unsuccessful because of the difficulties in reliably defining the transitions between components. An attempt is being made to solve this problem by developing a computer program which derives component times from curves fit to respiratory data.

PROGRESS: Two studies comparing respiration during the viewing of stress and neutral films have been conducted. A draft report describing these studies has been completed.

MAJOR FINDINGS: The results show no effect of the stress films on total breath times and no correlation between breath times and subjective ratings of mood. However, expiration times significantly increase and pause times significantly decrease during the stress films.

Analytical Ability in Verbal Functioning (NIMH, 737-2)

P. Oltman
H. Witkin

PURPOSE: It has been repeatedly found that measures of the cognitive style dimension of field dependence-independence are related at low and mostly nonsignificant levels to vocabulary and other related verbal abilities. However, it seems possible that there do exist some aspects of verbal functioning, especially those requiring breaking up of contexts, which would relate to field dependence.

PROGRESS: In order to explore this possibility, 18 verbal tasks of varying types, including word knowledge, fluency, context-breaking, etc., were administered to a large group of college students, along with two traditional measures of field dependence. In addition, written and spoken free-response verbal samples were obtained from a smaller group, all selected to be high on verbal ability, and subdivided into field-dependent and field-independent subgroups. The resulting data will be analyzed in an attempt to uncover a cluster of verbal tasks which share common features with the field-dependence measures.

Laterality Study (NIMH, 737-2) P. Oltman

PURPOSE: Increasing lateralization of functioning ("leftness vs. rightness") in various response modalities has been observed over the course of development. Whether or not various types of lateralization (hand, eye, visual field, ear, etc.) are related remains an unanswered question. In order to explore this domain, a battery of tasks known to show lateral differences will be administered and intercorrelations computed. Some preliminary results suggest that extent of lateralization is related to the field dependence-independence cognitive style, relations will also be examined between the lateral differences on each of the tasks and field dependence. It may be, for example, that there exist several unrelated clusters of laterality tasks and that cognitive style is related to certain of the clusters and not others.

PROGRESS: Pilot testing and final selection of the battery is in progress.

Studies of Young Adults with Uneven Cognitive Development (NIMH, 737-2)

P. Oltman
H. Witkin

PURPOSE: Factor analytic studies of the Wechsler Intelligence Scales have revealed that the subtest scores fall into three main factors: verbal comprehension, analytic, and attention-concentration. This study focuses on the first two of these. The first factor involves what is usually meant by verbal ability: vocabulary, information, and comprehension. The analytic factor is actually quite similar to the field dependence-independence cognitive style. The fact that verbal and analytic or field-dependence tasks are generally uncorrelated implies that while some individuals would be relatively consistent in normative level on these two factors, others would be discrepant in varying degrees. To study some of the implications of evenness vs. unevenness in level of verbal and analytic performance, various groups can be composed, some of which show equally high, medium, or low performance in the two areas, and others which show discrepancies favoring either verbal or analytic performance.

PROGRESS: A group of college students with verbal ability within the upper quartile has been selected, and further selection within this group has yielded a subgroup equally high on analytic ability and a second, discrepant group, which is in the lower quartile on analytic performance. A wide variety of verbal behavior and other data has been sampled in these groups, including structured and free-response verbal tasks, interviews, vocational interest data, and college transcripts. Intensive study of these sources of data will provide evidence on the impact of verbal-analytic discrepancy on verbal functioning, on patterns of interests, and on various aspects of the individual's academic career.

Cognitive Patterning in Sighted Children with Bilateral Retinoblastoma (NIMH, 737-2)

H. Witkin
P. Oltman

PURPOSE: In previous research we found that children totally and adventitiously blind from retinoblastoma were markedly superior in analytical ability (that is, they were field independent), but not in verbal-comprehension ability or attention-concentration ability. The frequently cited cognitive superiority of children with retinoblastoma thus seems specific to the analytical domain. This phenomenon is of particular interest against the fact that retinoblastoma is caused by a specific autosomal gene defect, although at this moment the findings obtained may as well be explained on an experiential basis. This study is the first in a series intended to pursue further the possible roles of genetic and learning factors in the cognitive patterns observed in retinoblastoma children.

PROGRESS: The subjects of this study are children with bilateral retinoblastoma, in whom vision was saved, so that while having the gene defect, they also had the same opportunities for cognitive learning as normal sighted children. Such cases are quite rare and the search for them is being continued as each child is located he is assessed on an extensive battery of cognitive tests. (Study being conducted in collaboration with the Columbia Presbyterian Eye Institute.)

Cognitive Style as a Factor in Dyadic Interaction Under Conditions of Initial Conflict (NIMH, 737-2)

H. Witkin
D. Goodenough
P. Oltman

PURPOSE: Previous studies have shown that participants in an interaction (teacher-student, patient-therapist, peers) who are similar in cognitive style are more likely to form positive impressions of each other, even after relatively brief periods of time together, than participants with contrasting cognitive styles. In the studies which yielded this result, the underlying social ground rule was one of working together toward a common goal (to have the student learn, to make the patient feel better). It is reasonable to expect, however, that other ground rules might lead to other outcomes with regard to mutual attraction, reflecting the role of situational factors as moderator variables in influencing the effects of match or mismatch. In the present study the social ground rule was one of initial conflict between the partners in the dyadic interaction with the requirement that the conflict be resolved. Under these conditions it was predicted that, because relatively field-dependent persons are more accommodating to the point-of-view of another and less likely to express or even experience opposition or hostility to others, dyads with at least one field-dependent partner would be more likely to arrive at a resolution of the conflict and, therefore, end up with positive feelings toward each other than dyads with two field-independent partners.

PROGRESS: Each of 40 subjects was required to spend a period of time in discussion, on one occasion with someone similar in cognitive style, and on another occasion with someone dissimilar. The task was to work out a basis for agreement about an issue on which, in pretesting, they were found to disagree. All sessions were videotaped, allowing analyses of both verbal and nonverbal communication processes. One result of the data analyses already completed is that dyads consisting of two field-independent partners showed significantly greater negative feelings toward each other (as expressed in a post-interaction questionnaire) than dyads consisting of a field-dependent and a field-independent partner or two field-dependent partners. A method of classifying modes of conflict resolution has been devised and is now being applied to the interaction dialogue. It is anticipated that the "ended-in-disagreement" category of this classification scheme will occur most often in dyads with two field-independent partners, whereas the "change-in-opinion" category will more often occur in the other two kinds of dyads.

Cross-cultural Studies of the Development of Psychological Differentiation (NIMH, 737-2)

H. Witkin
P. Oltman

PURPOSE: The objective of this study is to assess the role of socialization experiences in the development of psychological differentiation.

PROGRESS: Children from two small villages in each of three countries (Holland, Italy, Mexico) were subjects in this study. The pair of villages in each country was selected, on the basis of sociological data, as presenting a contrasting picture with regard to the particular child-rearing practices found in previous studies to influence the development of differentiation. Fifty 10-year-olds (25 boys and 25 girls) and 50 13-year-olds (25 boys and 25 girls) were studied in each country by a broad battery of cognitive tests and the figure-drawing test; their mothers were interviewed as well.

MAJOR FINDINGS: In each of the three countries differences in mean scores on tests of differentiation between children in the two villages were in the expected direction and were consistently significant. Findings on age and sex differences were in accord with results of many previous studies. Again in keeping with past results, measures of differentiation were significantly intercorrelated, but showed little relation to verbal measures. A paper on the results of this study was completed and submitted for publication.

Modes of Representation of Presleep Experiences in Dreams (NIMH, 737-2)

H. Witkin
D. Goodenough

PURPOSE: This study seeks to investigate the processes by which waking thoughts, images and feelings find transformed representation (in symbols, metaphors, etc.) in the language of the dream. Of particular interest is the relation between cognitive style and individual differences in mode of representation of presleep experiences in dreams.

PROGRESS: Data gathering for this study has already been completed, and data analyses are under way. The subjects, half of whom were field dependent and half field independent, came to sleep in the laboratory, where their EEG and eye movement were monitored, on five occasions, a week apart. The first of these was a practice session, and on two of the remaining four occasions subjects viewed a stress film immediately prior to sleep and on the other two occasions viewed a neutral film. Subjects were awakened from each rapid-eye-movement period and asked to report anything that had been going through their minds prior to awakening. The large body of dreams collected in this way, following upon the various presleep stimuli used, provide the data now being analyzed for exploring the relation between presleep experiences and dreams as a function of cognitive style.

Teacher-Student Interaction, Student Achievement and Mutual Attraction between Teacher and Students as a Function of Match or Mismatch in Cognitive Style (NIMH, 737-2)

H. Witkin
D. Goodenough
P. Oltman
C. A. Moore
F. McDonald

PURPOSE: One aim of this study is to confirm the finding of past investigations that teachers and students similar in cognitive style tend to evaluate each other more positively than teachers or students who are mismatched. Another purpose is to determine whether students learn more and are more motivated to learn when taught by teachers similar to them in cognitive style. Still another purpose is to identify the processes of interaction through which match or mismatch influence teachers' and students' feelings toward each other and student achievement. A further purpose is to identify features of teaching style and learning style which are a function of cognitive style. A final purpose is to identify the ways in which teachers, whatever their own cognitive style may be, adapt their teaching methods according to the cognitive style of the student being taught. It is anticipated that the data gained in pursuing the last purpose will contribute to the development of teaching methods specifically appropriate to students of different cognitive styles.

PROGRESS: A study is now under way in which teachers (both men and women) selected as extremely field dependent or field independent are teaching groups of four students, two field dependent (one boy and one girl) and two field independent (one boy and one girl), for four 50-minute sessions.. A curriculum on the "Rise and Fall of Mayan Civilization" was especially prepared for the study, both to suit its minicourse format and to allow for expression of differences in interests and in hypothesized modes of teaching and learning.

A variety of instruments and procedures is being used to assess student achievement as well as mutual attraction between teachers and students and among students. All class sessions are being audiotaped and videotaped and several parameters of teacher behavior and teacher-student interaction are being coded directly by coders stationed behind a one-way mirror. Before teaching the experimental classes, each teacher has a week of practice in which he or she goes through the entire procedure with a practice group of students. The study will use 24 teachers, 96 experimental students and 96 practice students as subjects.

P. Oltman
H. Witkin
D. Owen

Psychological Differentiation as Related to Antisocial Behavior, Impulsivity, and Endocrine Functioning in a Prison Population (NIMH, 738-1)

PURPOSE: Past studies have shown that psychopathology takes different forms in more differentiated and less differentiated persons when they suffer psychological breakdowns. Taking criminal behavior as also reflecting in some degree lapses in control, this study seeks to determine the forms criminal behavior assumes in more differentiated and less differentiated persons charged with criminal offenses. Also being examined in this study is the relation, in a group of men, between extent of field dependence and testosterone level. This relation is of particular interest against the repeated finding that not only do men tend to be more field dependent than women, but within the sex those who score in the more masculine end of masculinity-feminity scales tend to be more field independent. Finally, the study is examining the relation between testosterone level and aggression.

PROGRESS: 195 prisoners from the psychiatric prison service of a large New York City municipal hospital were the subjects in this study. Using all records available on each prisoner (the arrest report, prison chart, probation report) classification of crimes and methods of carrying out crimes were made, according to several hypotheses regarding the relation, first between psychological differentiation and forms of antisocial behavior, and, second between aggression and endocrine functioning. Similarly, classifications have been made of biographical data gathered through questionnaires administered to these prisoners. Finally, a scheme for assessing nature and effectiveness of controls was developed and applied to the figure drawings made by the prisoners. These data, together with the results of a battery of tests of field dependence, which were given to all the prisoners, are now being analyzed.

Relationship of Height to Criteria of Antisocial Behavior (NIMH, 738-1) — D. Owen

PURPOSE: It has been suggested that taller adolescents are more likely to be identified as delinquent than their shorter peers. The possibility has also been raised that convicted criminals who are taller than average may receive longer sentences than similarly charged men who are shorter. It is important to determine the extent, if any, to which added height influences the criteria of antisocial behavior, such as arrest record, conviction, or length of sentence. Therefore, an attempt is being made to assess the relationship between height and the existence of an arrest record. Also, an attempt is being made to assess the extent to which similar criminal cases, after a guilty verdict has been obtained, are disposed of differently as a function of the height of the criminal defendant.

PROGRESS: From a starting population of 31,455 consecutive male births during 1944-1947 in Copenhagen, Denmark, the heights of 28,588 individuals still living have been obtained. All civil and penal violations charged against any of these men and the dispositions of their cases have been collected from the police. A coding scheme to accommodate these crimes and resulting criminal sanctions has been devised, and the data have been coded. A computer tape file for the criminality data is being created, after which analyses will begin.

Genetic Factors in Field Dependence and in Aggressive Behavior: Studies in Sex Chromosome Anomalies
(NIMH, 738-1)

H. Witkin
D. Goodenough
D. Owen

PURPOSE: This study is attempting to identify noninstitutionalized men with two kinds of sex chromosome aberrations, one with an extra Y chromosome (XYYs), the other with an extra X chromosome (XXYs). The data obtained with the groups thus identified, along with the data from comparison groups of normal XY men, will be used for two purposes: first, to check a model of the role of the sex chromosomes in field dependence, and second, to check the hypothesis recently put forward that an extra Y chromosome contributes to the development of aggressive behavior. Since groups institutionalized for antisocial behavior, which are of course highly biased populations, have been used in almost all the work done on the aggression hypothesis to date, what is especially needed are studies of noninstitutionalized groups.

PROGRESS: A case-finding study is now being completed in Copenhagen, Denmark, where social records are particularly well suited both for identifying subject populations and for obtaining evidence on antisocial behavior. To this point, all males (31,380) born in Copenhagen between 1941 and 1945 have been identified and the present location of all those now alive has been determined. Using height data available for these men at the draft boards, the group in the top 15 per cent for height have been selected for sex chromosome determinations. Ninety per cent of these determinations have already been made. Controls are now being selected for the groups of XYYs, XXYs and translocations that have been identified. The extensive battery of psychological and physical tests to be given to the sex-chromosome anomalies and their controls has already been developed and the testing of these cases is to begin in the fall.

Sex Chromosome Anomalies in a Male Prison Population
(NIMH, 738-1)

H. Witkin
P. Oltman
D. Owen

PURPOSE: The proposal has been made that men with an extra Y chromosome are overrepresented in prisoner groups, compared to the general population, suggesting that the extra Y chromosome is associated with aggressive behavior. At the moment this "aggression hypothesis" is hardly proven. The data on hand bearing on this problem have often come from studies inadequately controlled and suffering a number of selection biases.

PROGRESS: Nine hundred eighty-three consecutive admissions to the prison ward of the psychiatric service of a large New York City municipal hospital were karyotyped. The small number of XYYs (only three) identified in this group is not beyond the frequency to be expected in the general population. This finding is inconsistent with the aggression hypothesis. The difference between the XYY frequency in our group and the higher frequency of XYYs found in some other mental-penal groups that have been examined may be related to differences in admission policy. It is noteworthy that four XXYs were also identified in our prison group, again a frequency not different from chance. The similar frequency of XYYs and XXYs is consistent with the results of several other recent studies. A paper on the results of this study is now being prepared for publication.

Reference Measures for Cognitive and Personality Factors (ONR, 739-1) — H. Harman

PURPOSE: Under the sponsorship of the Office of Naval Research, ETS is conducting research to identify and to present measures of well-established factors by updating the previous Kit of Reference Tests for Cognitive Factors and by extending it to the noncognitive domain.

PROGRESS: Some 28 personality factors have been identified from a search of the literature, and more than 1400 items were written to measure them. These items were put in 30 booklets (with repetitions) and administered in as many sessions to more than 4000 recruits at the Naval Training Center, San Diego. By means of a complex design, it was possible to limit the testing time for any individual to no more than 320 items while still enabling the determination of relationships among all 88 subscales designed to mark the 28 factors. Analysis and interpretation of results are now under way.

As a result of a presentation of the experimental design of the study made for the Society of Multivariate Experimental Psychologists an offer came for a tryout of our personality items at the University of Oregon. The total set was put in two booklets (following the same principle of random assignment of items) and administered to a sample of approximately 250 undergraduates. These students had already taken two other personality inventories. It will thus be possible to look at the structure of our own instrument in comparison with the structure of two other instruments.

Two technical reports are in process: Toward the Establishment of Noncognitive Factors Through Literature Search and Interpretation; and Cognitive Factors: Some Recent Literature. Early results on the personality factors of the current project are being presented at the August 1973 meetings of the American Psychological Association in a Symposium: Toward the Development of More Comprehensive Sets of Personality Measures.

Perceptual Factors in Cognition and Personality (ONR, 741-10)

S. Messick
J. French

PURPOSE: This study investigates factor analytically the generality in cognitive and personality domains of the perceptual factors of speed and flexibility of closure.

PROGRESS: Experimental tests of verbal and cognitive closure and flexibility in reasoning and problem solving were developed and included in a test battery with selected personality and reference ability measures. The test battery was administered to more than 500 cadets at the U. S. Naval School of Aviation Medicine, Pensacola, Florida. A factor analysis of intercorrelations among 35 scores on perceptual and cognitive tasks has produced about 14 factors, which were rotated to oblique simple structure. Twenty scores from personality measures were also projected into the space by extension methods. Examination of the results indicated factors interpretable in terms of speed and flexibility of perceptual closure, of verbal closure, and of cognitive closure. Intercorrelations among the primary factors were factor analyzed to produce four second-order factors. A hierarchical transformation was performed to place the first- and second-order factors in the same orthogonal framework and to obtain the loadings of the tests on the second-order factors. The second-order factors appear to reflect verbal comprehension, general reasoning, figural articulation (i.e., facility in structuring and differentiating figural stimuli), and symbolic articulation. A report has been completed and will appear as a Research Memorandum.

Automated Hypothesis Testing (ONR, 746-3; 746-4; 772-21; NSF, 816-1) — F. Lord, M. Stocking

PURPOSE: To construct and implement an algorithm to test standard and nonstandard hypotheses and automatically compute standard errors for large samples.

PROGRESS: The program has been tested on about two dozen assorted multivariate problems with success.

MAJOR FINDINGS: An algorithm has been constructed and is reported in RB-72-42, revised edition. This report will be submitted to a journal for publication. A computer program with instructions for its use has been written up (RM-73-7).

Estimating Item Parameters (ONR, 746-3; 746-4; NSF, 816-1)

F. Lord
M. Wingersky

PURPOSE: To find a practical method to estimate the parameters of the characteristic curve of each item in a test.

PROGRESS: A method for estimating ability and item parameters using information provided by omitted responses has been developed and programmed.

MAJOR FINDINGS: "Estimation of Latent Ability and Item Parameters When There Are Omitted Responses" (RB-73-37) describes the method and gives theoretical and empirical justification for it. The program, documented in RM-73-2, is available for general use.

Strong True-Score Theory (ONR, 746-3; 746-4)

F. Lord
M. Wingersky
M. Stocking

PURPOSE: The purpose of this project is to find a mathematical model for the details of the relation between true score and observed score, fitting a wide range of ordinary test data, so that one can estimate or predict in advance the mental test performance of groups of examinees in ways not previously possible.

PROGRESS: Improvements were made to the computer program for estimating true-score distributions and to the program for estimating the bivariate distribution of test scores from the marginals. In addition the bivariate program was modified to compute the relative efficiency of two tests.

Dimensions of Perceived Item Similarity (USPHS, 752-2)

S. Messick

PURPOSE: To ascertain the number and nature of dimensions of perceived similarity among personality statements (in terms of similarities in likelihood of endorsement) in an effort to appraise the relative contributions of perceived content consistencies and perceived desirability consistencies in determining judgments of endorsement frequency.

PROGRESS: Twenty-three MMPI items were selected to span six dimensions previously obtained in analyses of MMPI responses. All possible pairs of these 23 statements were presented to 150 college students, who were asked to rate the similarity of the members of each pair with respect to their probability of endorsement. The 150 x 150 matrix of sums of cross-products among individuals was computed and factored according to the individual differences model of multidimensional scaling. Four dimensions of viewpoint about the similarity of these items appeared, and separate multidimensional spaces were derived for each viewpoint. These perceptual spaces were rotated orthogonally, and the dimensions were interpreted. One important distinction reflected in most of these spaces is the separation of desirable from undesirable item content. A report is in preparation. (Study done in collaboration with Dr. D. N. Jackson of the University of Western Ontario.)

Leveling-Sharpening as a Cognitive Style
(USPHS, 754-7)

F. Damarin
V. McKenna
S. Messick

PURPOSE: This research has three major purposes and several minor ones. The first purpose is the replication and extension of earlier findings by Gardner, Jackson, and Messick on the cognitive style of leveling-sharpening. The tests which are involved in this research include many psychophysical measures. A second purpose is to discover whether the leveling-sharpening trait or other psychological traits can serve as the parameters of the psychophysical law, as Damarin has suggested they should. Since many leveling-sharpening measures are similar in their formal structure both to measures of figural aftereffects and to measures of retroactive inhibition, a third purpose of the research is to discover whether "levelers" also show larger FAE's and more retroactive inhibition as McKenna has hypothesized.

PROGRESS: Forty-seven objective tests, including putative measures of leveling-sharpening, field independence, category width, neuroticism, figural aftereffects, retroactive inhibition, verbal satiation, rote memory, memory span, meaningful memory, and psychophysical judgment, and a variety of questionnaires, including most of the known response styles, were administered to 147 freshman men and women at the College of William and Mary in Williamsburg, Va.

MAJOR FINDINGS: In an earlier study, now completed, cognitive style measures showed an impressive ability to predict individual differences in the recall of photographs of human faces, thus supporting assertions to this effect by Witkin (cf. RB-63-21).

A later study developed evidence that several different types of leveling-sharpening may exist. One of these is reasonably similar to the dimension that Gardner and his associates interpreted as the tendency to assimilate perceptions to memories.

The same research yielded evidence that individual parameters of the psychophysical function may correlate over different types of judgments, giving some support to Damarin's "trait law" hypothesis. A more refined analysis of these data is now under way.

McKenna's factor analysis of memory tests suggests that there are memory styles associated with a preference for errors of omission vs. commission. A report on this work is available (λB-68-28).

Personality Organization in Cognitive and Social Processes (USPHS, 754-7) — S. Messick

PURPOSE: This project attempted (a) to clarify in a series of multivariate studies the nature of several cognitive styles and to investigate their interrelations with measures of intellectual ability, attitude and belief structures, affect expression and control, academic and creative achievement, and personality; (b) to analyze sex differences in cognitive and personality structure; (c) to assess the stability of cognitive styles and the correlates of reliable changes; and (d) to evaluate the role of cognitive styles in interpersonal and social functioning.

PROGRESS: (a) Categorizing Styles and Cognitive Structure (with Nathan Kogan). This study attempts to clarify two basic questions: (1) Is there a general dimension of breadth of categorization yielding individual consistencies in categorizing behavior across varied stimulus domains and measurement procedures; and (2) what are the cognitive, intellectual, and personality correlates of stylistic consistencies in categorization? Group-administered measures of verbal and quantitative category-width judgments, object sorting, cognitive complexity, preference for complexity, and extremity of judgment were administered to approximately 160 college students, along with several other cognitive, intellectual, and personality variables. Factor analyses were performed separately for male and female subjects. The results indicate that categorizing consistencies exist, but that they are multidimensional and their organization is somewhat different for the two sexes. Females appear to display more generality across categorizing tasks than do males: several dimensions relevant to categorizing emerged for both sexes, but they appeared to be somewhat more differentiated and task specific for males. Personality measures were found to be implicated with categorizing and judgmental consistencies in females, whereas they were more directly associated with intellectual abilities in males. Although the emergence of several categorizing factors in this study indicates the operation of differential stylistic tendencies in this domain, the factors were at the same time found to be correlated with each other, suggesting some generality at a second-order level. Factor analyses of the intercorrelations among the primary factors produced three second-order factors for males and two for females. In both sexes, one of the second-order factors reflected general categorizing consistencies that cut across the stimulus and process distinctions embodied in the various first-order factors.

(b) Cognitive Complexity: Dimensions of Role Constructs (with Nathan Kogan). In conjunction with the study of categorizing styles described above, an attempt was made to assess certain aspects of cognitive complexity using multidimensional scaling techniques. Much recent work on cognitive complexity has used a procedure developed by George Kelly for assessing personal constructs--the Role Construct Repertory Test (Rep Test), in which the subject is presented with several triads of figures significant in interpersonal relations, such as mother, self, and best friend. The subject is then asked to indicate which two members of the triad are most alike and to write down both the basis for their similarity and the reason why the third member is different. The number of different constructs listed has been used as a measure of cognitive complexity.

The 160 subjects in the categorizing study described above were administered a modified form of the Rep Test, and they were also asked to judge on a 10-point scale the degree of similarity or difference between all possible pairs of the 15 role figures. Cross-products matrices for these similarity ratings were computed separately for males and females and were analyzed according to the individual differences model of multidimensional scaling. Six dimensions of viewpoint about stimulus similarity were obtained for the females and seven for the males. The individual scores on these dimensions were then correlated with the cognitive and personality measures included in the categorizing study in an attempt to ascertain some of the properties of individuals holding each viewpoint. In this case, meaningful patterns of significant correlations were obtained between the viewpoint scores and such measures as acquiescence, impulsiveness, rigidity, unconventionality, preference for complexity, category width, and conceptual differentiation.

Two composite stimulus spaces were also derived, one to represent the similarity judgments of the males on the average and one to represent the similarity judgments of the females on the average. For the females, four dimensions were required to represent the average similarity judgments among 15 stimuli, whereas for the males three dimensions appeared adequate. The same three dimensions appeared in both spaces, but the fourth female dimension provided a distinction in sex-role relations (they described similarities between "mother" and "sister" on a separate dimension from similarities between "father" and "brother") that was not emphasized by the males. One report is available and another is in preparation.

(c) Cognitive Interference and Flexible Control. The term "constricted vs. flexible control" refers to consistent modes of reacting to interfering and contradictory cues. This cognitive control has been most frequently assessed in terms of interference scores on the Stroop Color-Word Test. In the present study, a group-administered version of the Stroop Color-Word Interference Test was developed, as well as three other group procedures designed to measure susceptibility to cognitive interference in the face of differing amounts of competing response and at various levels of cognitive involvement. An attempt was made in developing these procedures to vary the ease of separating competing cues in the interference conditions, so as to permit an empirical separation of two related processes thought to be important in withstanding cognitive interference effects: (1) facility in restricting attention to the relevant aspects of the stimulus and the response and (2) facility in actively inhibiting the competing response and other disrupting tendencies. The four tasks were administered to approximately 90 college males, along with the original individually-administered Stroop Color-Word procedure and several other cognitive and personality measures. Both control and interference sections of all five Stroop-type tests were found to correlate highly with each other (median r of .53), but specific processes involved in each task were differentiated in a factor analysis, as were the cognitive and personality correlates of the factors. A factor analysis of intercorrelations among nine first-order factors uncovered two second-order dimensions, which appeared to reflect (1) the ability to withstand disruption in serially organized tasks and (2) the ability to isolate and

identify the appropriate aspects of the stimulus and response. A regressed interference score was also obtained for each of the five procedures to reflect that part of the interference performance which was independent of the control performance. These regressed scores possessed substantial reliability, and although they were not related to each other, they exhibited different patterns of significant correlations with personality measures.

(d) Leveling-Sharpening and Assimilation Tendencies (with Fred L. Damarin and Virgil McKenna). The cognitive style of leveling-sharpening has been postulated to account for consistent individual patterns of performance in psychophysical tasks and in the relative differentiation of free association responses. These cognitive consistencies have been attributed to individual differences in the activity of the memory process of assimilation. Levelers assimilate in the sense that they tend to merge perceived objects or stimulus events with similar but not identical objects recalled from previous experience.

In an effort to clarify possible multiple determinants of leveling-sharpening consistencies, the present study attempted to specify the concepts of assimilation and contrast in many different ways as a basis for developing a variety of new experimental tests.

One procedure that was highlighted finds its prototype in the Schematizing Test, which has been used as a criterion measure of leveling-sharpening in other laboratories. The Schematizing Test provides two scores: a lag score measures the subject's tendency to underestimate change in a series of squares that gradually increase in size, and an accuracy score measures his ability to estimate magnitudes accurately within a fixed range of stimulus values.

By also including tasks with different types of stimuli and different judgmental conditions, the present study focused upon the construct validity of alternative interpretations of leveling-sharpening and assimilation-contrast. A battery of several experimental psychophysical judgment tasks was administered to 150 college students, along with measures of field-independence, category width, figural aftereffects, retroactive inhibition, verbal satiation, memory abilities, and various personality characteristics. A factor analysis produced 17 factors, which were rotated to oblique simple structure. An analysis of the intercorrelations among the first-order factors produced four second-order dimensions. One of the factors could very plausibly be interpreted as a measure of assimilation or dedifferentiation of memory traces over time, but the lag score from the Schematizing Test was not related to it. Another factor appeared to represent uncriticalness in judgment, a general tendency to judge similar stimuli as the "same" as opposed to "different," but again the lag score was unrelated. However, the schematizing score was found to be associated with one of the second-order dimensions, which could be interpreted in terms of leveling-sharpening. One report has been completed (RB-68-28) and another is in preparation.

(e) Tolerance for Unrealistic Experiences (with Norman Cliff). This dimension, also called tolerance for instability, refers to the differential

willingness to accept perceptions at variance with conventional experience. The concept was introduced originally to account for data linking form-boundedness and form-lability on the Rorschach with responses to unusual kinds of perceptual experiences, such as those arising in apparent movement and aniseikonic distortion.

The present study attempted to appraise the reliability of several measures of apparent movement and aniseikonic distortion, and to evaluate the intercorrelations among them for evidence of stylistic consistencies. Measures of aniseikonic distortion and perceived range of apparent movement were obtained from a sample of 50 undergraduates. Moderate reliabilities were obtained for certain measures of aniseikonic distortion. High reliabilities were achieved for movement threshold and simultaneity threshold scores in the apparent movement procedure, but the two thresholds correlated positively with each other (.46). Correlations between aniseikonic measures and the movement and simultaneity measures were small, but some consistencies appeared. There was a slight tendency for distortion measures to correlate more strongly with movement thresholds than with simultaneity thresholds, and some distortion measures correlated in the same direction with both thresholds. Taken together, these findings suggest that scores for range of apparent movement reflect, in addition to a willingness to perceive the illusion of movement, a possible assimilation phenomenon in which subjects vary in their tendency to stick with a particular judgment (either alternation or movement) once they have made it.

(f) Scanning and Focusing (with Harold Schiffman, Duke University, and Diran Dermen). The dimension of scanning-focusing refers to individual differences in the extent of spontaneous attention deployment. Extensive scanners tend to search and peruse a field of interest, surveying both relevant and irrelevant or miscellaneous properties. The present study attempted to assess scanning and focusing consistencies in a wide variety of tasks in an effort to obtain some convergence of correlations to clarify the interpretation of the construct. These procedures included perceptual judgment tasks (such as size and distance estimation), perceptual speed tasks, and perceptual search tasks (in which the subject must locate in a larger field stimuli of a particular class). Since scanning propensities may also be reflected in the manner in which internal memory fields are surveyed, measures of fluency of ideational production and remoteness of response in word association were also included. A specific attempt was made to differentiate between two possible types of scanning--scanning for signal detection and scanning for information seeking. This was done using perceptual search tasks in which the subject was required to locate stimuli (signals) embedded in meaningfully organized visual fields, e.g., locating faces camouflaged in pictorial scenes. Upon completion of the search task, the stimulus materials were removed, and the subject was then asked specific questions about the content of the scenes. Subjects who incidentally take in information about the field in the process of scanning could thus be differentiated from those whose concern is apparently limited to detecting the signal. Measures were also included for facility in detecting signal stimuli in unorganized fields, such as locating four-letter words in arrays of letters. In view of a theoretical and empirical link between scanning and

isolation of affect, several tasks yielding measures of affect expression and control were also included, such as early childhood memories, the Holtzman Inkblot Technique (HIT), and a picture preferences procedure that assesses consistent tendencies to like, dislike, or be indifferent to a wide variety of photographs. These procedures were administered to about 100 male and 120 female college students, along with measures of potentially relevant cognitive styles (particularly field independence and category width), intellectual abilities, and personality.

Factor analyses were performed separately for males and females. Although rotations are not yet complete, some striking relations are discernible.

In males, for example, a factor reflecting incidental knowledge of the content of the field was found to be relatively independent of a factor reflecting facility in locating signal stimuli in unorganized arrays; skill in finding faces hidden in organized fields was associated with both dimensions. In females, a factor emerged linking size estimation with form appropriateness and the absence of shading on the HIT, the latter being a traditional Rorschach indicator of conflict about the acceptance of affect. In both analyses, several cognitive dimensions emerged that displayed substantial correlates with inkblot measures and personality scales.

Another feature of this study was an analysis of errors or "wrong" scores, which was undertaken to investigate the possibility that scanners might be differentiated from nonscanners by the type and number of errors they make. Factor analyses of these "wrong" scores revealed several processes that apparently lead to errors on these tasks. These factors include haste and carelessness, intellectual deficiency, premature perceptual closure, impulsiveness, and susceptibility to distracting and embedding contexts.

(g) Cognitive Styles, Defenses, and Eye Movements (with Lester Luborsky, University of Pennsylvania Medical School, Harold Schiffman, Duke University, and Diran Dermen; partly supported by an NIMH grant to Dr. Luborsky).

Extreme scanning has been found by others to be marginally related to ratings of isolation, projection, and generalized delay on the Rorschach (with isolation possibly related to an "information seeking" type of scanning involved in an obsessive concern with exactness to offset doubt and uncertainty, and with projection possibly related to a "signal detection" type of scanning involved in a paranoid concern with accuracy to offset suspicion and distrust). On the other hand, minimal scanning or focusing has likewise been considered to reflect a concern for accuracy and a preference for viewing the world in a narrow discriminating way. The present study attempts to clarify some of these issues by relating measures of scanning and other cognitive styles (leveling-sharpening and breadth of categorizing) to projective measures of defenses and to characteristic patterns of eye movements in viewing affective and nonaffective stimuli. Heart rate measures of arousal were also obtained under basal control conditions and while viewing the affective and nonaffective pictures.

These procedures have been administered to a small sample of about 40 college students. Several eye-movement scores (such as number of fixations, mean time per fixation, mean track length, and scatter) were obtained for the viewing of each of ten photographs. Heart rate measures were also obtained separately for each photograph presentation. The internal consistency reliabilities of these scores over all ten photographs and over subcategories of photographs reflecting sexual, aggressive, and neutral content have been assessed. Many of the scores proved to be substantially or highly consistent across all of the pictures, but a few were reliable for only one or two of the three subcategories. Such differential reliability was to be expected, since differences in stimulus pattern and meaning should have had more of an impact on some of these scores than on others. The interrelations among eye movement, heart rate, defense, cognitive style, and personality measures have been appraised and a report is in preparation.

(h) Style and Persuasibility (with Fred L. Damarin). This study investigates the interrelations among response styles, cognitive styles relevant to susceptibility to field forces (such as field independence and flexible control), and susceptibility to persuasion and opinion change. A battery of such measures was administered to approximately 90 male college students.

A factor analysis of attitude shift scores, response styles, and selected cognitive and personality measures produced ten factors. Although rotation to oblique simple structure has not yet been completed, preliminary results reveal the presence of four factors clearly related to persuasibility, two of them relatively specific to particular attitude sources and two of them more general in that consistencies are displayed in response to several types of attitude source.

(i) Sex Differences in Cognitive and Personality Structure (with Diran Dermen). This study was designed to bear upon several interrelated questions about the nature and generality of cognitive styles. Of foremost concern were (a) the determination of the number and nature of primary dimensions necessary to account for previously observed stylistic consistencies in cognition; (b) an appraisal of the interrelations of these stylistic dimensions with measures of intellectual ability, attitude and belief structures, affect expression and control, academic and creative achievement, and personality; and (c) an evaluation of sex differences in the psychological organization of cognition and personality.

A battery of experimental measures that required over 20 hours of testing time was administered to more than 700 high school juniors and seniors over a two-month period. The sample included approximately equal numbers of males and females, as well as representatives of college preparatory, commercial, vocational-industrial, and art curricula. High school students were sought for this study because of the extreme difficulty of finding comparable samples of males and females at the college level, i.e., samples where either explicit- or self-selection had not occurred differentially by sex on variables relevant

to the sex comparison, not only with respect to obviously relevant variables (like ability patterns and interests) but with respect to more subtle ones as well (like maternal education level). The procedures have been scored not only for the specified dimensions under study, but also for various other response properties (such as number of incorrect or unacceptable answers, number of items omitted, and specific response sets) that would help to characterize the test-taking behavior more completely and might provide important controls for clarifying the basis of some of the obtained relations. Since many of the techniques entailed the evaluation of free responses, the scoring itself was a major undertaking that took more than a year to complete. Reliabilities have been assessed on the total sample and separately for male and female subsamples. They have, with few exceptions, proved satisfactory. Further analyses are in progress.

(j) The Stability of Cognitive Styles and the Correlates of Change (with Diran Dermen). This study attempts to assess the stability of cognitive style measures over a one-year period and to determine the cognitive and personality correlates of any reliable changes. A sample of 135 high school seniors, who had been extensively tested a year earlier as part of the previous study, were subsequently retested on selected measures of field independence, element articulation, form articulation, conceptual differentiation and compartmentalization, analytic vs. relational categorizing, susceptibility to cognitive interference, and scanning.

Test-retest reliabilities over the one-year interval will be determined, and differences in corresponding means and variances between the two testing sessions will be evaluated. Base-free measures of change will be calculated for each test, in order to partition the difference in true scores between the two testing sessions into two components: one completely dependent upon initial position in the first session and one reflecting that part of the second performance that is completely independent of the previous performance. The reliabilities of these base-free change scores will be evaluated prior to computing their correlations with each other and with the cognitive and personality measures of the first battery.

In addition, a secondary focus of this study dealt with the malleability of categorizing consistencies. Subjects were required to sort the objects in the object-sorting measure of conceptual differentiation first under standard instructions and then in a new and different way. Facility in changing a natural category system could thereby be assessed, as well as the manner in which the change was brought about.

These indices of flexibility and style in changing conceptual categories will be correlated with cognitive style measures from the concurrent battery and with cognitive and personality measures from the previous battery.

Retest Changes on Personality Scales
(USPHS, 767-2)

L. Stricker

PURPOSE: A ubiquitous tendency exists for scores on self-report personality scales to change over time, and in the direction of improved "adjustment," even though no experience intervenes that would be expected to produce such changes. This study seeks to clarify the meaning of these score changes by investigating the nature of individual differences in them. The study will determine the generality of individual differences in score changes on diverse personality scales, as well as the relationships of these individual differences to score changes and initial scores on cognitive and personality measures (not of the self-report type) selected for their relevance in testing hypotheses about the causes of these trends.

PROGRESS: Needed cognitive and personality measures were adapted or developed and administered with the Sixteen Personality Factors Test, the California Psychological Inventory, and the Personality Research Form in a pilot study. The data were analyzed and measures selected for use in the main study. The data for the main study have been collected and the statistical analysis is under way.

Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma (ARPA, 769-1)

N. Kogan
M. Hermann

PURPOSE: The study has two major aims. The first is to examine the effect of being a representative of a group on behavior in the Prisoner's Dilemma (PD). To date in most research on the PD, subjects have acted in their own behalf, the interaction with the opponent being of major concern. What happens, however, if the opponents in the PD must not only be concerned about each other's behavior but also about how well they are representing a reference group? The second aim of the study is to explore the effects of decision style on responses in the PD. By decision style is meant a subject's way of approaching a decision-making task--e.g., his flexibility, his risk-taking dispositions.

PROGRESS: A preliminary report of the study was presented at the 1969 EPA meetings. A final report has been issued (RB-72-45), and a version for publication is currently in preparation.

MAJOR FINDINGS: The results suggest that whether or not a person is acting in his own behalf or representing a group in the PD will affect how he behaves. Representatives as opposed to individuals tend to be more cooperative to start with in the PD, particularly the highly committed representative. A relationship was established between personality and behavior in the PD. This success may stem from using the dyad instead of the individual as the unit of analysis and from examining the effects of personality across trials.

Negotiation Behavior under Conditions of Observation and Consultation (ARPA, 769-1) — N. Kogan

PURPOSE: In an earlier study (RB-67-51), leader-delegate dyads came to joint decisions, then were separated into all-leader and all-delegate negotiating groups with instructions to resolve conflicts of interest and achieve consensus. Differences were observed between the two group types, but these were not of great magnitude. We suggested that the differences might have been attenuated by the fact that leaders and delegates could exert no direct influence upon each other after the dyads were separated. In the present experiment, leaders and delegates negotiated in each other's presence. In half of the groups, leaders negotiated while delegates served as observers and consultants; in the other half, the roles were reversed. Differential negotiation outcomes were expected, given the inequality of power and influence between the leader and delegate roles.

PROGRESS: A report (RB-69-55) has appeared, and a revised version has been published. (Negotiation Constraints in the Risk-Taking Domain: Effects of Being Observed by Partners of Higher or Lower Status. Journal of Personality and Social Psychology, 1972, 23, 143-156.

MAJOR FINDINGS: Predicted differences which failed to appear in the previous study (RB-67-51) were observed in the present research. Delegates negotiating in the presence of their leaders (relative to leaders negotiating in the presence of their delegates) were more likely to deadlock, to seek the opportunity for private consultation with their dyadic partners, and to consume more time both in consultation and actual group discussion. An additional finding of interest was the greater dissatisfaction with group decisions expressed by those in the role of observer than by those in the role of negotiator. (Project conducted in collaboration with Dr. H. Lamm and Miss G. Trommsdorff of the Institute for Social Sciences, University of Mannheim, Germany.)

Study of Aptitude-Achievement Relationships (770-1) J. Carroll

In connection with participation in a conference on the aptitude-achievement distinction that was organized by Dr. Henry Kaiser (University of California, Berkeley) and Dr. Donald R. Green (CTB/McGraw-Hill), Carmel, California, January 1973, Mr. Carroll developed a mathematical model of aptitude-achievement relationships as they apply over multiple school grade levels. This model was an extension of his "model of school learning" as published originally in Teachers College Record, 1963, 64, 723-733. The model postulates that individual differences in aptitude are reflected in different rates of learning, and that these in turn are reflected in measurement achievement levels at different school grade levels. However, measured aptitude is regarded as at least partly or potentially a function of time because of the increased maturity, experience, and "test wiseness" of the examinee. The model makes predictions concerning regressions of achievement on aptitude at different school grades or points of time in the school career of the student. It also makes predictions concerning the means and variances of aptitude and achievement measures. The model was applied to data provided by Dr. George Burket of CTB/McGraw-Hill concerning batteries of aptitude and achievement tests that are vertically equated for grades 1-12 and that were administered to about 175,000 pupils in grades 1-12 on a nationwide scale. The relatively good fit of the model to the obtained data tends to confirm the validity of the model. A preliminary report of this work was given at the January 1973 conference and also at the AERA convention, New Orleans, February 1973 under the title "The Importance of the Time Factor in Learning." The final version of the report of this work will be published in a book edited by D. R. Green of CTB/McGraw-Hill, probably in late 1973. Further application of the model to longitudinal data available at ETS is planned.

Cooperative Literature Tests (770-32) P. Diederich

PURPOSE: The purpose of this project is to discover types of difficulties in adult books that baffle and discourage many students at some point near age 14 when they must make a transition from juvenile to adult books if their interest in reading books on their own is to continue.

PROGRESS: Cooperative Tests and Services had already initiated the development of tests on the 20 literary works that are most widely taught in grades 9-12. Editing, tryout, norming, and statistical analysis of over 3,000 items on these works afforded an opportunity to explore in depth the types of questions that less than 50 per cent answered correctly in 211 schools and school districts in 42 states and the District of Columbia. The tests are now ready for publication except the one on Ivanhoe, for which too few cases were found. The next step will be the preparation of a monograph on the types of difficulties in this sample of adult books that were revealed by item analysis.

Application of the Lognormal Model of Word Frequency Distribution (770-33)

J. Carroll

PURPOSE: In previous work, it was demonstrated that the lognormal distribution yields good fit to word frequency distributions when account is taken of the bias in drawing finite samples from a theoretical population. The lognormal theory makes possible the estimation of the type-token function and the vocabulary size underlying a given sample. The research aims to explore the application of the lognormal model to various types of samples in order to aid in the characterization of individual differences in verbal ability and the description of samples of verbal material. It also aims to develop a psycholinguistic and mathematical rationale to explain and predict the lognormal distribution or some approximation thereof.

PROGRESS: Work has continued on the application of the lognormal theory to further data that have become available from various sources, and on the preparation of a monograph to report the results.

The theoretical orientation of this project has led to explorations of several topics connected with word frequency, its measurement, and its correlates:

(1) The data from an experiment described in the report for 1971-72 (RB-72-10) have been incorporated in a published article, "Word Frequency and Age of Acquisition as Determiners of Picture-Naming Latency," Quarterly Journal of Experimental Psychology, 1973, 25, 85-95, by John B. Carroll and Margaret N. White. In multiple-regression analysis of picture-naming latencies from an experiment modelled on Oldfield and Wingfield's (1965), with 94 stimuli and 37 adult subjects, two word frequency measures had insignificant beta weights, while two measures estimating age at which the word was learned had highly significant weights. Objects whose names were learned early were named faster. This result may have important implications for the interpretation of studies using word frequency as a critical variable. It is suggested that word retrieval may be a one-stage process that depends upon the age at which a word was learned.

(2) To obtain data for the further evaluation of age-of-acquisition as a word attribute in studies of verbal behavior, learning, and memory, estimates were secured from 62 undergraduates (35 males, 27 females) of the age at which they believed thay had learned each of 220 picturable nouns (divided into two lists assigned randomly to halves of the sample), according to a 9-point scale. Reliabilities of these ratings were about .98. For comparative purposes, word frequency values for the words were secured from three large word-count studies or, where necessary, from subjective estimates made by 20 adults. Use of these and other variables as predictors of the previously obtained picture-naming latencies (see above) yielded results supporting the previous finding that age of acquisition is a more relevant predictor than word frequency. Some word frequency indices tend to reflect age of acquisition, but when this

influence is minimized word frequency makes little contribution to prediction. These results were reported in Research Bulletin (RB-72-58), and an article based on this RB will appear in the October 1973 issue of the Journal of Verbal Learning and Verbal Behavior.

(3) Data collection has continued for a further experiment to study picture-naming latencies. One part of this experiment is essentially a replication of the study described in paragraph (1) above, but with 50 words selected so that word-frequency and age-of-acquisition measurements are as orthogonal as possible, and with three trials given to every subject. Subjects are also given several cognitive ability tests, so that we may explore relations with picture-naming latencies. The results of this part of the experiment have been analyzed; they completely confirm the previous finding that age-of-acquisition is a more potent variable than word frequency in picture naming. An interesting finding also is what may be termed a "Rip Van Winkle" effect whereby words long dormant in memory suddenly acquire high strengths in a second trial and then lose part of that strength in a third trial. In the subsequent phases of the experiment, not yet completed, an effort will be made to see whether latencies are affected if the subject has prior knowledge of either the word-frequency or the age-of-acquisition classification of the word whose picture he will be naming. These studies are regarded as important in investigating the properties of "truly long-term memory" (TLTM).

(4) The study of word-retrieval processes has led to the consideration of other choice-reaction phenomena, in particular to the study of (musical) absolute pitch judgments. At the September 1972 APA convention Mr. Carroll gave a preliminary report of findings based on two subjects claiming absolute pitch ability, one musically trained subject not claiming absolute pitch ability, and one subject who claimed to have trained himself in this ability (P. T. Brady, Journal of the Acoustical Society of America, 1970, 48, 883-887). It was clear from these preliminary results that persons with absolute pitch ability can transmit at least five bits of information in absolute pitch judgments (as opposed to the two to three bits for most other sensory modalities), and that latency of judgment is a function of the number of bits being transmitted. During the remainder of 1972-73, data were collected for 5 further subjects, 2 claiming absolute pitch ability and 3 not so claiming. The data are being analyzed in preparation for the publication of a report.

(5) Further work has been undertaken on developing a mathematical rationale for the asymptotic lognormal word frequency distribution as reported in preliminary form in RB-69-90. At the suggestion of several statisticians, the generating function being used is that of the Beta distribution. It is hoped that the results will be sufficiently promising to complete a further publication on this topic.

Factors in Comprehension of Verbal Material (770-33)

J. Carroll
R. Freedle

PURPOSE: This project was envisaged as a program of separate studies to determine various parameters in the comprehension of discourse and in the learning of facts, concepts, and principles from such discourse. Particular attention would be paid to the role of lexicon ad syntactical structure. Although some of the work would be done with native speakers of English responding to oral and written discourse in English, studies would also be concerned with the comprehension of discourse in the learning of a second language in order to examine the development of comprehension in detail. Also, comprehension ability would be studied developmentally by conducting experiments and tests with individuals sampled from different age levels, particularly the years of schooling.

PROGRESS: During 1972-73, the major new development under this project was the working out of a preliminary model of a "performance grammar" for a subset of English sentences. This model was motivated by the belief that before comprehension processes can be properly studied, it is necessary to develop a theory of how sentences (and utterances in general) are produced from semantic elements assumed to be cognitively prior to their actual encoding into surface structures. It was felt also that rules of the type proposed in transformational generative grammar would be inadequate to account for sentence generation as a behavioral event or for the acquisition of language in the young child. Rather, rules interpretable as behavioral tendencies or learned habits would be necessary. The present model contains a series of contingency rules whereby various kinds of sentences (active, passive, declarative, interrogative, "querying," positive, "challenging," and negative) involving two noun phrases related to a transitive verb may be produced in a left-to-right manner from given semantic elements and variables. The model has been translated into an operating computer program. Further work will involve expansion of the grammar to a greater variety of sentence types both in English and in other languages (to compare the rules involved in speaking those languages). It will also involve testing the assumptions of the model by psycholinguistic experimentation, and application to the description and explication of child language acquisition. The model can also be used to make predictions concerning sentence comprehension processes. An article describing this work has been accepted for publication in the International Review of Applied Linguistics.

Other work related to this project has been concerned with various publication activities. The proceedings of the Research Workshop on Language Comprehension that was held at Durham, N. C. in the Spring of 1971 under the sponsorship of the NAS-NRC/NAE Committee on Basic Research on Education were published as a book, under the editorship of Mr. Freedle and Mr. Carroll, Language Comprehension and the Acquisition of Knowledge, Washington: V. H. Winston & Sons, September 1972.

As a member of the Committee on Literacy appointed by the National Academy of Education in response to the late James E. Allen, Jr.'s request to make a report on research and practice in the field of reading, Mr. Carroll has continued his work on editing the final report of the Committee. This report, which will include a number of position papers contributed by leading experts in the field, will be published as a book, Toward a Literate Society, under the editorship of John B. Carroll and Jeanne S. Chall by McGraw-Hill Inc., in 1974.

The study of people's perceptions of nouns as mass nouns or count nouns, reported in an ETS Research Memorandum (RM-72-2), was accepted for publication in a Festschrift volume for Professor Archibald Hill of the University of Texas. It is understood that this volume is in press at Mouton & Company, The Hague.

The study performed by Professor Donald Spearritt of the University of Sydney during his tenure as Visiting Research Psychologist at ETS from September 1971 to January 1972, and reported in Research Bulletin RB-72-4, was accepted for publication in the Reading Research Quarterly.

Psycholinguistic Theory and Models (770-34) R. Freedle

PURPOSE: The purpose of this project is to investigate language from its earliest inception through adulthood so as to isolate important psychosocial factors which contribute to and modify language development and language use. Particular areas of investigation will include the following topics: (1) the analysis of sequential vocalization behavior of mother and infant as a function of social and situational setting from age 3 months to 2 years; (2) substantive models which will account for the regularities observed in language development from its prelinguistic phase (3 months) to its more formal stage (2 years); (3) the relationships of psychosocial factors on language comprehension, production, imitation, and communication; (4) information-processing models to account for the psychosocial factors on language use; and (5) story comprehension and recall as a function of psychosocial factors as well as a semantic model to describe the internal structure of stories presented for recall.

PROGRESS: In collaboration with M. Lewis the following reports are available: (1) for prelinguistic vocalization in interaction with his mother see RB-72-22; (2) for models of sequential vocalization see RB-71-34; (3) for psychosocial studies see RB-73-18, with William S. Hall; (4) for information-processing approaches to language usage as a function of various psychosocial factors see Freedle and Hall's "An information-processing approach to some problems in developmental sociolinguistics," presented at the biennial meetings of the Society for the Study of Behavioral Development, Ann Arbor, Michigan, 1973; also see Hall and Freedle's "A sociolinguistic account of the influence of peers on the language of black and white children," also presented at the Ann Arbor meetings, 1973. For other information-processing models see RB-73-9 and RB-73-19; (5) for studies of prose comprehension and topic identification as well as message production see Freedle's "Language users as fallible information-processors: Implications for measuring and modeling comprehension" in Freedle and Carroll's Language Comprehension and the Acquisition of Knowledge. Washington, D. C.: V. H. Winston & Sons, 1972.

MAJOR FINDINGS: A Markov model describing six states of the vocalization interactions between mother and infant appears to be moderately accurate in reproducing the sequential states over a two-hour observation period. Further, measures of the transitional probabilities from this analysis appear to be correlated with linguistic measures obtained on the same children at age two years. A full report of the latter findings is in progress. Situational settings significantly affect the sequential flow of vocalizations between mother and infant; furthermore, at different SES levels it appears that the infants may be "learning" different vocalization patterns in interacting with their mothers. Also significant sex differences at this early age can already be noted in vocalization behavior.

An information-processing model for describing the responses made to sentences presented in either standard or so-called nonstandard English indicates that blacks and whites perform equally well when rate of

improvement measures for standard English (from age 8 to 10) are employed. In addition, memory storage and retrieval parameters of the information-processing model indicate that the two races are equal in their response to standard English; they differ in which is called a "dialect-preference" parameter which measures the degree to which each population chooses to code the stimulus sentences in either standard or nonstandard form--the latter parameter is characterized as a nonability factor and is seen instead to reflect a number of conflicting cultural pressures to use one dialect over another in a variety of communication settings.

Other information-processing approaches to sentence recall indicated that a serial processing mechanism was favored over a parallel processing one; steps in the information-processing decision tree implicated syntactic and semantic strategies in sentence comprehension for children of mean age 4 1/2 years.

Personality Research Form: Factor Structure and Response Style Involvement (771-7) — L. Stricker

PURPOSE: The Personality Research Form (PRF) is a comparatively new personality inventory. This instrument was constructed by specially developed procedures which were designed to insure that its scales were very homogeneous, relatively independent of each other, and minimally influenced by acquiescence and SD response styles. The aims of this study were (a) to explore the factor structure of the PRF and (b) to examine the inventory's relations with response styles.

PROGRESS: The inventory and standard measures of acquiescence, social desirability, and defensiveness response styles were administered to a sample of high school students. The PRF content scales were factor analyzed and the loadings of the stylistic measures on these factors were estimated.

MAJOR FINDINGS: Six oblique factors, identified as conscientiousness, hostility, ascendance, dependence, imagination, and carefreeness, were found. The stylistic measures' estimated loadings on these factors were scattered and moderate. A report is available (RB-73-23).

Response Styles and 16 PF Higher-Order Factors (771-7) L. Stricker

PURPOSE: There have been some suggestions that second-order factors on self-report personality scales may reflect response styles, such as acquiescence and social desirability, rather than such content traits as extroversion, anxiety, and neuroticism. This study is intended to test such response-style hypotheses by examining the relationships between standard measures of these variables and second-order factors on Cattell's Sixteen Personality Factors Questionnaire (16 PF).

PROGRESS: The inventory and standard measures of acquiescence, social desirability, and defensiveness response styles were administered to a sample of high school students. Four second-order and one third-order factors were extracted from the 16 PF. The factors were rotated and their relationships to the response style measures were determined by extending the response style measures into the factor matrices.

MAJOR FINDINGS: Three of the four second-order factors corresponded to well-established 16 PF second-order factors. Each kind of response style measure predominantly loaded a different second-order factor. Social desirability scales defined a novel factor somewhat resembling an anxiety factor, and acquiescence and defensiveness measures loaded but did not define independence and anxiety factors, respectively. None of the various kinds of response style indexes consistently loaded an extroversion factor, and none loaded the third-order factor. This factor was not similar to previously reported 16 PF third-order factors and its nature was unclear. A report is available (RB-73-8).

Convergent and Discriminant Validation of the DPI (771-24) — D. Jackson

PURPOSE: To investigate the convergent and discriminant validity of the Differential Personality Inventory (DPI), an instrument designed to measure certain areas of psychopathology.

PROGRESS AND MAJOR FINDINGS: Roommate ratings and self-ratings by 370 college subjects were used as comparison measures. For the total sample, all DPI content scales correlated significantly ($p < .001$) with relevant self and peer ratings. Results from a factor analysis revealed a clear organization of 11 trait specific factors with high loadings for the relevant trait measured by each of the three methods. This study has been published in the Journal of Clinical Psychology, 1973, 24, 214-219. (Study done in collaboration with Dr. Kenneth A. Carlson of the Ontario, Canada, Department of Correction.)

The Effects of Detoxification on Psychopathology for Alcoholics as Measured by the Differential Personality Inventory (771-24)

D. Jackson

PURPOSE: To assess changes in psychopathology at various stages of alcoholic detoxification and in a treatment program following detoxification.

PROGRESS: The Differential Personality Inventory (DPI) was administered to two groups of consecutively admitted male alcoholics to a state hospital, the first a Detoxification Group, and the second, a group already detoxified.

MAJOR FINDINGS: Of the 27 personality scales, there was a significant reduction in 19 for the Detoxification Group. Scales showing mean decrements included those reflecting Insomnia, Headache Proneness, Broodiness, Cynicism, Depression, Disorganization of Thinking, Familial Discord, Feelings of Unreality, Hypochondriasis, Ideas of Persecution, Irritability, Mood Fluctuation, Neurotic Disorganization, Panic Reaction, Perceptual Distortion, Rebelliousness, Self-Depreciation, Somatic Complaints and Infrequency. Scales reflecting personality and character disorders, such as Hostility, Shallow Affect, and Socially Deviant Attitudes, did not change significantly during detoxification. Other results indicate additional reduction in psychopathology following detoxification, again largely in neurotic and psychotic symptom areas. Data from a DPI validity scale indicated that even patients suffering from acute alcoholic toxicity could answer purposefully. Test-retest data reveal that although systematic mean changes in both groups occur, individuals show a marked tendency to retain their relative position within a group following detoxification and treatment. An article based on these findings was completed in collaboration with Helmut Hoffmann and Paul C. Nelson of Minnesota State Hospital, Willmar, Minnesota, and accepted for publication in the Journal of Clinical Psychology.

A Four-Dimensional Interpretation of Risk Taking (771-24) D. Jackson

PURPOSE: To investigate the dimensionality of risk taking.

PROGRESS: A preliminary conceptual analysis led to the hypothesis that the risk-taking construct was comprised of at least four major facets: Monetary, Physical, Social, and Ethical risk taking. Four personality scales, one for each facet, were developed, together with three additional methods for measuring each facet: Self-Rating, Situational Dilemmas, and Vocational Preference. These sixteen measures were incorporated into a multitrait-multimethod matrix, with certain additional personality scales.

MAJOR FINDINGS: Substantial evidence for convergent and discriminant validity was found, both from correlational and factor analytic data, providing support for the four-facet model proposed. Even though four independent factors were obtained, however, a second order factor analysis yielded strong evidence for an underlying dimension of generalized risk taking. Implications for research were discussed. An article based on these findings, written in collaboration with Neil J. Vidmar and Larry Hourany of the University of Western Ontario, has been published in the Journal of Personality.

Judged Frequency of Endorsement and Frequency of Occurrence Scale Values and Dispersions for MMPI Items (771-24)

D. Jackson
S. Messick

PURPOSE: To distinguish perceived frequency of occurrence and perceived frequency of endorsement of a personality trait, and to prepare sets of scale values useful in personality research.

PROGRESS AND MAJOR FINDINGS: The scaling method of successive intervals was applied to nine-point rating data based on judgments of frequency of occurrence and on judgments of frequency of endorsement for all items in the booklet form of the MMPI, obtained from 111 college students, divided randomly into groups of 55 and 56 for their respective sets of judgments. Judged frequency of occurrence and judged frequency of endorsement scale values and discriminal dispersions for these items are reported. Evidence was obtained for their stability and certain applications noted and discussed. An article has been completed and accepted for publication.

A Model for Inferential Accuracy (771-24) D. Jackson

PURPOSE: To develop a theoretical model to account for individual differences in the process of trait inference, which involves the use of limited information in forming a judgment about another person.

PROGRESS AND MAJOR FINDINGS: Inferential accuracy is distinguished from other forms of accuracy in social perception, including differential and stereotype accuracy, and defined in terms of a person's ability, given limited information about a target person, correctly to judge other pertinent characteristics about that person. A model and accompanying paradigm are proposed for studying this process. Given the observation that judgments about the joint probability of traits tend to be highly stable across groups of judges, it is possible to obtain scale values for a set of personality characteristics for each target. Each judge's ratings of this target may be plotted with respect to these ranked scale values in such a way that the degree to which his own judgments parallel the consensual patterning of trait inferences may be determined. Two parameters, slope and intercept, determine a straight line summarizing this relationship. These parameters are hypothesized to correspond to the judge's sensitivity to trait inferential information, and his threshold for employing more or less remote inferential data. It is hypothesized that these two parameters will account for a major portion of the individual difference variance in tasks involving inferential accuracy. The evidence to date, although limited, suggests that while sensitivity is general across dimensions, threshold is more specific. Whereas sensitivity is potentially linked to accurate knowledge of lay personality theory, threshold may be related to a process similar to assimilative projection. An article describing this model has been published in the Canadian Psychologist.

Studies in the Factorial Interpretation of Multitrait-Multimethod Matrices in Personality Assessment (771-24) D. Jackson

PURPOSE: To distinguish method and trait variance with factorial methods and to apply these methods to the investigation of the heteromethod validation of personality questionnaires.

PROGRESS: After consultation with Professor Ledyard R Tucker of the University of Illinois, a method of evaluating factor reliability with implicit factor scores has been developed and a flexible computer program to accomplish this has been written. This method has been adapted to permit the sequential factoring of heteromethod-heterotrait matrices in a manner which facilitates the separation of trait and method variance.

MAJOR FINDINGS: A number of matrices involving the same set of traits measured by different methods have been evaluated. Results indicate that it is quite possible to obtain multimethod convergent and discriminant validity without the use of Procrustes-type rotation of axes or alternative fitting of data to preconceived targets.

Data are also being gathered on a national sample for the evaluation of common factors between the domains of vocational interests and personality. The use of the above model will permit the analysis of these common factors without the intrusion of factors specific to either vocational interests or personality.

Intercultural Attitude Study (772-11)

H. Gulliksen
D. Gulliksen

PURPOSE: This study is using balanced incomplete block designs, complete paired comparisons, and successive intervals designs, with factor analysis and analysis of variance techniques, in comparing attitudes and preferences of different national groups. Eight different questionnaires relating to work, goals, aims, and activities have been given to respondents in Italy, Poland, Sweden, and the United States (Pennsylvania).

Scale values were obtained and analysis of variance on votes for items or on category choices. Factor analysis of covariances and varimax rotation with means added proved to be a valuable technique for showing the structure of group attitudes.

MAJOR FINDINGS: Consistency of response was indicated by relatively few total circular triads (TCTs) over $p = .025$. Females showed significantly fewer TCTs than males at the $F_{.95}$ level. Positive rewards in work were: creativity, self-direction, change and variety, learning and advancement, recognition, pay and security. The quality of personal and family life took precedence over an interesting job. Other interests outside of work were still lower in scale values. This study was issued as an Office of Naval Research Technical Report in April 1971. The revised manuscript has been issued as a monograph by Multivariate Behavioral Research (No. 72-3).

Learning (772-11; NICHHD, 786-9) H. Gulliksen

PURPOSE: A study of learning curves in cats was conducted in an attempt to replicate learning curves in divided brain animals.

PROGRESS: Learning data on normal and split-brain cats has been analyzed by a curve fitting program, developed by James Ramsay, which maximizes the likelihood function for the Audley-Jonckheere learning model by the method of conjugate gradients. A learning model earlier proposed by Gulliksen is a special case of the Audley-Jonckheere model. In the current analysis, values were obtained for the parameters of the model: Rho_0, Alpha, Beta, Gamma 1 and Gamma 2. The parameter Rho_0 represents the initial probability of a correct response, Alpha represents the amount by which a correct response increments the correct habit strength, and Gamma 1 minus Alpha represents the amount by which a correct response increments the incorrect habit strength. Similarly, Beta represents the amount by which an incorrect response increments the correct habit strength, and Gamma 2 minus Beta represents the amount by which an incorrect response increments the incorrect habit strength.

MAJOR FINDINGS: The values of Rho_0 varied from .28 to .62. A plot of the "left brain" vs. the "right brain" values of Rho_0 did not show agreement of right with left brain values. The values of Alpha were generally small, and the values of Gamma 1 minus Alpha were zero or very small. The values of Beta were smaller than those for Alpha, and values for Gamma 2 minus Beta were very small or zero. Plots of Beta and Gamma 2 minus Beta did not show agreement of left and right brain values. These parameters varied a great deal.

Data from the computer analyses of the learning problems, following the Audley-Jonckheere model and from the conventional learning measures were analyzed to obtain correlation coefficients and other indices of agreement between values for left and right brain learning. Three indices of agreement between first paw and second paw learning were computed:

$$(1) \qquad r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$$

$$(2) \qquad R_1 = \frac{2\Sigma xy}{\Sigma x^2 + \Sigma y^2}$$

$$(3) \qquad R_2 = \frac{2\Sigma xy - \frac{1}{2N}(\Sigma X - \Sigma Y)^2}{\Sigma x^2 + \Sigma y^2 + \frac{1}{2N}(\Sigma X - \Sigma Y)^2},$$

where

$\underline{r}$ is the usual Pearson correlation coefficient. It equals unity if $Y = aX + b$ or if $y = ax$, it equals zero if $\Sigma xy = 0$, and it equals -1.0 if $y = -ax$.

$\underline{R}_1$ depends only on deviation scores. It is unity if $X_i - \bar{X} = Y_i - \bar{Y}$ or $x_i = y_i$ for all i, and equals zero if $\Sigma xy = 0$, and equals -1.0 if $x_i = -y_i$ for all i.

$\underline{R}_2$ is an index which equals 1.0 if $X_i = Y_i$ for all i, and equals zero if the variance of $(x + y)$ equals the variance of $(X - Y)$, and equals -1.0 if $X_i = -Y_i$ for all i.

Since it is theorized that learning measures in the split brain animals would be equal for first and second paw performance on the same problem, the values of R_2 should be large if the measures were equal. Where R_2 is small, even though $\underline{r}$ is large, it indicates that though the correlation may be high, the values of X and Y are far from equal. R_1 expresses the relationships between deviations from the mean.

The values of r, R_1 and R_2 were generally near zero indicating lack of agreement between first and second paw performance.

The revised manuscript was issued as an Office of Naval Research Technical report and as ETS Research Bulletin 73-46. (Study done in collaboration with Dr. T. Voneida of Case Western Reserve University.)

SAT Delta Study (772-21) F. Lord

PURPOSE: To find out how effective are the present SAT and the five experimental tests designed by the Division of Analytical Studies and Services at various ability levels. Also to investigate the properties of various hypothetical SAT tests.

PROGRESS: Item parameters were estimated for 85 SAT Math Items. Using these parameters, information functions were computed for the five subtests as well as for the regular SAT. Information functions for various hypothetical SAT Math tests were computed using regular scoring and optimal scoring.

MAJOR FINDINGS: The five experimental tests were similar to each other but differed from the regular SAT by being slightly more discriminating below the 75th percentile of ability and somewhat less discriminating between the 90th and the 99th percentiles. Both the regular SAT and the experimental tests provide poor measurement for the bottom sixth of the students. An unpeaked SAT provides better discriminating power than would a peaked SAT over a wide range of abilities.

A 60-item SAT Math test that would be equally discriminating throughout the range of 275 to 790 gives only four-tenths as much information as the regular SAT in the 400 to 620 range and is about as effective throughout the total range as a 10-item peaked SAT would be at the ability level for which it was designed. For low ability levels, optimal scoring would approximately double the information obtained.

Power Scores Estimated by Item Characteristic Curves (772-21; NSF, 816-1) F. Lord

PURPOSE: A group of students were improperly timed in the regular administration of the SAT Verbal test. The purpose of this study was to estimate by item characteristic curve theory the score each individual would have obtained if he had been allowed to finish the test.

PROGRESS: Item characteristic curve parameters were estimated and power scores computed.

MAJOR FINDINGS: RB-72-48 discusses the method for estimating power scores, giving supportive empirical results, and applies the procedure to 21 students who were mistimed.

The Relative Efficiency of Two Tests (772-21; NSF, 816-1) F. Lord

PURPOSE: The Birnbaum formula for the relative efficiency of two tests, using the ratio of two information functions, requires the estimation of the item characteristic curve parameters. The purpose of this project was to develop a practical formula that may be easily computed.

PROGRESS: A new formula was developed that expresses relative efficiency solely in terms of the standard errors of measurement and the frequency distributions of true scores. This formula may be approximated using the frequency distributions of raw scores.

MAJOR FINDINGS: "The Relative Efficiency of Two Tests as a Function of Ability Level" (RB-73-41) derives the new formula and gives a numerical illustration comparing this method with the method that uses the item characteristic curve parameters.

A Discussion of Certain Reliability Formulas (772-63) W. Kristof

PURPOSE: The accuracies of four different formulas for reliability estimation are compared under certain distributional assumptions.

PROGRESS: The work will be presented at the 81st Annual Convention, 1973, of APA and will be contained in the Proceedings.

MAJOR FINDINGS: There is a uniformly most accurate reliability estimator. Its general use is recommended.

On Accuracy in Reliability Estimation (772-63) W. Kristof

PURPOSE: This study in parametric test theory deals with the statistics of reliability estimation when scores on two parts of a test follow a binormal distribution with equal (case 1) or unequal (case 2) expectations. In each case biased maximum-likelihood estimators of reliability are obtained and converted into unbiased estimators. Expressions for the accuracies of the estimators as functions of the reliability parameter and sample size are sought.

PROGRESS: The work is contained in RB-73-24.

MAJOR FINDINGS: Second moments of the estimators are obtained and utilized in calculating mean square errors of estimation as a measure of accuracy. A rank order of four estimators is established. There is a uniformly best estimator. Tables of absolute and relative accuracies are provided for various reliability parameters and sample sizes.

W. Kristof

On the Statistics of Reliability Estimation Using Parallel Test Forms (772-63)

PURPOSE: Inferences about the reliability of a given test require repeated measurements in one form or another on a sample of subjects. Two approaches to data collection are common: (a) one obtains multiple measurements using basically the same test whose reliability is the quantity of interest; (b) one obtains multiple measurements using comparable parts of the test whose reliability is the quantity of interest. In the second case the reliability of the component parts is stepped up to give the reliability of the total test. This procedure is not required in the first case. One might therefore assume that case (b) should lead to a statistical theory more complicated than that based on case (a). However, the opposite is true. Too little work with emphasis on statistics has been presented for case (a). Hence it is intended to close a gap in existing mental test theory by developing the statistical reliability theory for case (a).

PROGRESS: A number of pertinent results have been obtained. A Research Bulletin is in preparation.

MAJOR FINDINGS: Suppose that two test forms with bivariate normal score distribution are available. Then the results include the following: (1) The sampling distribution of the maximum-likelihood estimator $\hat{\rho}$ of reliability parameter ρ ; (2) The expectation of $\hat{\rho}$; (3) Bounds for the bias of $\hat{\rho}$ as a function of ρ and sample size N ; (4) The minimal bias of $\hat{\rho}$ as a function of ρ and N . Obtaining these and a number of similar results requires extensive handling of hypergeometric and other transcendental functions. Methods of advanced calculus are needed.

Analysis of Covariance Structures (772-67) M. Browne

PURPOSE: Methods are to be developed for estimating parameters in covariance structure models. Properties of the estimators are to be investigated.

PROGRESS: A Research Bulletin has been issued (RB-73-1).

MAJOR FINDINGS: Under certain general conditions on the covariance structure model, generalized least squares estimators are asymptotically efficient. The limiting distribution of a corresponding quadratic form in residuals is chi-square. Maximum likelihood estimators belong to the family of generalized least squares estimators. When the covariance structure is linear in the parameters to be estimated, some generalized least squares estimates involve far less computation than maximum likelihood estimates. Computational procedures have been developed.

Rotation of Factor Matrices (772-67)

M. Browne
M. van Thillo

PURPOSE: Effective algorithms are to be developed for rotating a factor matrix to simple structure. The problem is one of optimization subject to equality constraints.

PROGRESS: Constraints on the transformation matrix in both orthogonal and oblique rotation have been imposed by means of reparameterizations. Programs have been prepared for minimizing the Crawford-Ferguson family of criteria using the Fletcher and Polak-Ribière optimization procedures. A research bulletin describing the methods (RB-73-22) and a research memorandum describing the programs (RM-73-11) have been issued.

The Generalization Function in Probability Learning Experiments with Infinitely Many Stimuli and Responses (772-69)

M. Levine

PURPOSE: To discover a procedure for estimating the generalization function in those probability learning experiments in which both stimuli and responses are points on a continuum like length. To study the dependence of the generalization function upon the distribution of reinforcements. To develop methods suitable for studying individual differences in generalization.

PROGRESS: A procedure has been discovered for computing the generalization function directly from individual learning data. The procedure has been applied and found to be very precise. The common translation invariance assumption which had been supported by less precise and sensitive methods was found to be false in both the author's data and previously published data. However, it was discovered that all generalization functions satisfied a functional equation almost as simple as the translation invariance equation. The new equation permits one to analyze generalization functions into two components: one which is invariate over distributions of reinforcements and one which can be deduced (using Weber's law) from the distribution of reinforcements. Although consistent individual differences were observed, all subjects in all conditions conformed to the new equation.

Geometric Interpretations of Magnitude Estimation and Other Direct Psychophysical Measurement Procedures (772-69)

M. Levine

PURPOSE: To develop a simple, intuitive but quantitative theory capable of integrating the diverse experimental findings arising in the study of direct psychophysical measurement. To account for the data without the ad hoc and implausible hypotheses that have been used by many earlier theories.

PROGRESS: Using only some elementary projective geometry and some intuitions already in the psychometric and experimental psychology literature it has been possible to give a coherent account of many experimental findings. In particular, by assuming that perceptions can be represented as points in a multidimensional space and that orderings of unidimensional attributes are generated by processes like geometric projection one can derive the invariance of ratios of magnitude estimates with change of modulus, some range effects, the concurrence of certain families of magnitude estimation curves (such as those observed in the study of the size-weight illusion) and the regression effect. The theory also predicts a relationship between magnitude estimates and rating scales which seems to be present in published data.

MAJOR FINDINGS: A small number of classical psychological ideas can give a coherent, intuitive and precise account of the major quantitative findings of direct measurement procedures. Details can be found in RB-73-32 which is to appear as a chapter in Contemporary Developments in Mathematical Psychology edited by R. C. Atkinson, D. H. Krantz, R. D. Luce, and P. Suppes. The results on rating scales is being prepared for publication in Psychological Review.

Nonadditive Analogues of the Basic Mathematics of Additive Measurement (772-69)

M. Levine

PURPOSE: The most well-developed area of measurement theory is limited to applications like the measurement of length in which there are no important interactions and in which it is reasonable to assume that the measured objects combine additively. Many of the most important results in this area (additive measurement) depend on a remarkably small number of mathematical results. A type of nonadditive measurement (called affine measurement) is becoming important in modern experimental psychology and psychometrics. The main purpose of this study was to contribute to the development of affine measurement by finding nonadditive analogues of the basic mathematics of additive measurement.

PROGRESS AND MAJOR FINDINGS: Analogues of the key theorems of additive measurement have been discovered. In addition, a simple condition for experimentally testing the appropriateness of affine measurement and a procedure for reformulating nonadditive computations as additive measurement computations have been found. Details are given in RB-73-34. A full report has been submitted to the Journal of Mathematical Psychology.

Parameterizing Curves Transformable to the Same Shape (772-69)

M. Levine
D. Saxe

PURPOSE: Suppose a family of curves such as the item characteristic curves of a latent trait theory of mental test scores or the isosensitivity curves of signal detection theory can be transformed into curves with the same shape. Then each curve in the family can be identified by a pair of parameters: a mean-like location parameter and a standard deviation-like scale parameter. There are two special cases which are fairly well understood. Case I: The curves are known exactly but the transformation is unknown. In this case, algebraic methods from measurement theory (for example Levine's work on functional equations) can be used to compute the transformation and the parameters exactly. Case II: The curves are known only approximately but the transformation is known exactly. In this case statistical methods (for example Lawton's work on self modeling) can be used to estimate the parameters.

In many potential areas of application, the transformation is unknown and the curves are specified only approximately by data. The purpose of our studies is to develop and test algebraic methods from measurement theory for directly computing parameters from untransformed, empirical curves and to make the methods available in computer programs for general use.

PROGRESS AND MAJOR FINDINGS: Several algorithms have been discovered. They are being tested with Monte Carlo and mental test data. One algorithm involving a novel use of fourier series has been especially successful. When applied to the item test regression curves of the Scholastic Aptitude Test, item difficulties previously computed by Lord using maximum likelihood methods and Birnbaum's logistic model were reconstructed with great accuracy in spite of the fact that no parametric assumptions were made. These results will be reported at the Montreal Mathematical Psychology Meetings in August. A Research Bulletin describing the algorithm is now being reviewed.

Classification under the Model of Compound Symmetry (772-79) — L. Gleser

PURPOSE: We are interested in whether the problem of classifying p-variate normal random observations into one of k populations simplifies if it is known in advance that the mean vectors and covariance matrices of the k populations have the structure of compound symmetry first introduced by Votaw.

PROGRESS AND MAJOR FINDINGS: If the mean vectors and covariance matrices of the k p-variate normal populations are known, the form of all admissible classification procedures is obtained. The form of such procedures is identical to the form of the admissible classification procedures in a lower-dimensional subspace (of dimension $r < p$) of p-dimensional space.

When the covariance matrices and mean vectors of the k populations are not known (and have no special structure), but we have sample estimators of these quantities, the classification procedures of Anderson, of Wald, and the likelihood ratio procedures all require distributional results which are extremely complicated. Unfortunately, little simplification results when it is known that the covariance matrices and mean vectors in k populations have the compound symmetry structure. However, in the special case where the covariance matrices and mean vectors have a completely symmetric structure (the intraclass covariance model), enough simplification is obtained so that the needed distributional results are tractable (with the aid of a computer). Expansions for the probabilities of misclassification are given for this special case. The results are being prepared for publication in a Research Bulletin.

A Guide to Probability Theory and Applications (772-79; 772-81)

L. Gleser
I. Olkin

PURPOSE: The recent success of probability models in explaining and describing various social and behavioral phenomena has made it clear that the basic aspects of probability theory are now as important to the behavioral sciences as they have proved to be in the engineering and physical sciences. Probability models for learning, conformity, mobility, social grouping, information dispersal (rumors), infant-parent interaction, and so forth, have yielded important new insights. Yet, the number of psychologists, sociologists, and educational researchers with backgrounds in probability theory strong enough to enable them to both profit from these models and develop new probability models is quite small. Part of the reason for this scarcity of individuals with backgrounds in probability theory is that most books which discuss the basic concepts and models of probability theory in sufficient depth require a mathematical background in calculus, while probability books which do not require calculus tend to omit very important topics and models which are of use in the behavioral sciences. A new book designed to provide behavioral scientists with a solid understanding of probability concepts and models at an elementary mathematical level is most urgently needed at this time.

PROGRESS: The present book (co-authored with C. Derman) is designed to introduce the behavioral scientist to the essential features of probability theory at an elementary level (requiring only a background in college algebra). The material is divided into two parts: elementary concepts and probability models. The first part deals with the development of the basic probability calculus. There the pace is slow in order to permit those with a minimum of mathematical background to learn the material. The second part deals with material not generally available in texts and includes topics (such as multivariate models and Markov Chain models) which are of importance in understanding and developing probability models for the behavioral sciences. Serious examples from various disciplines have been used to exemplify the material. The book has been published by Holt, Rinehart, and Winston, Inc. (Library of Congress Catalog Card No. 72-88276) under the above title, and is now available.

L. Gleser
I. Olkin

Large-Sample Expansions for Null Distributions of Likelihood Criteria (772-79; 772-81)

PURPOSE: Many multivariate test statistics (such as the likelihood ratio test statistic for MANOVA) have null distributions whose moments are proportional to ratios of products of gamma functions. For any random variable W, $0 \leq W \leq 1$, whose moments have the above-mentioned form. Box proposed an asymptotic expansion for the cumulative distribution function of W, which provides an accurate method for determining the critical constants defining rejection regions for the multivariate tests mentioned above. Although the method is useful, the calculations needed to obtain the coefficients in each asymptotic expansion must be done ab initio, and almost always involve cumbersome algebraic manipulations. The purpose of the present study is to (i) simplify the needed computations, and (ii) find formulas for the coefficients of the expansion in special cases.

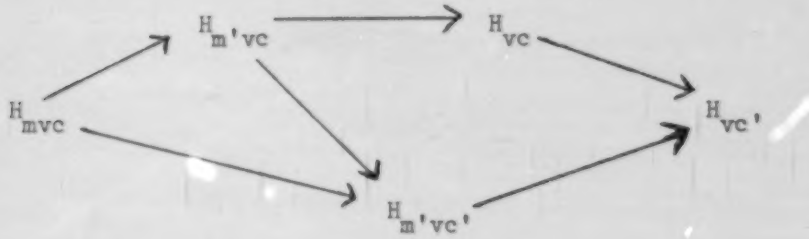
PROGRESS AND MAJOR FINDINGS: Simplified algorithms are given for calculating the coefficients of the asymptotic expansion in the general case. In a certain special case (which includes the null distributions of the likelihood criteria for MANOVA and for testing the independence among sets of variates), explicit formulas are derived for these coefficients. Finally, if $W_1, W_2, \ldots, W_k$ are k statistically independent random variables each of which has a cumulative distribution function which can be expanded in Box form, then it is shown that the cumulative distribution function of $W = \prod_{i=1}^{k} W_i$ also can be expanded in Box form. Formulas relating the coefficients of the expansion for the cumulative distribution function of W and the coefficients of the expansions of the cumulative distribution functions of $W_1, W_2, \ldots, W_k$ are obtained. Two examples of the use of the above methodology to find critical constants for multivariate tests of hypotheses are also provided. A Research Bulletin (RB-72-34) describes the main results of the study.

Multivariate Statistical Inference under Marginal Structure. I (772-79; 772-81)

L. Gleser
I. Olkin

PURPOSE: Suppose that we are utilizing k different psychological tests, which have one subtest T_0 in common, and which are designed to be parallel. The components of the g-th test can be represented as (T_0, T_g), where T_0 is the subtest common to all k tests, and T_g is the subtest peculiar to the g-th test. Of particular concern is the hypothesis that these tests are parallel with respect to the means and/or covariances.

PROGRESS AND MAJOR FINDINGS: A complete hierarchy of hypotheses in which the tests are "parallel" has been developed. For example, $H_{m'vc}$ is the hypothesis that the tests are parallel with respect to the means of the common test only, and with respect to the covariances of both tests. This hypothesis might be tested against H_{vc}, the hypothesis of parallelism with respect to the covariances. These hypotheses may be represented as follows:



in which an arrow indicates implication. The prime indicates equality for the common test only.

Maximum likelihood estimators under the various models (and under the assumption of normally distributed test scores) have been obtained, as well as the related likelihood ratio statistics. Approximate distributions of the likelihood ratio statistics are worked out, so that the tests can be applied. An example using SAT data is included. This material has appeared as a Research Bulletin (RB-72-40), and a condensed version is soon to appear in the British Journal of Mathematical and Statistical Psychology.

ultivariate Statistical Inference under Marginal
tructure. II (772-79; 772-81)

L. Gleser
I. Olkin

URPOSE: The present paper continues the study of statistically testing hether psychological tests are parallel. However, now a more hierrchical structure is permitted. Suppose three tests (perhaps in subequent years) are used, and suppose further that each test consists of hree subtests. We indicate a test by (A, B, C) to denote the parts. In he model considered, the tests have the following composition.

Test 1: (T_0, U_1, V_1)

Test 2: (T_0, U_1, V_2)

Test 3: (T_0, U_2, V_3)

hus, we see that subtest T_0 is common to all three tests, subtest U_1 is ommon to the first two tests, and the third part is peculiar to each indiidual test. The main concern is to test for parallelism with respect to he means and/or with respect to the covariances of thc three tests.

ROGRESS AND MAJOR FINDINGS: A likelihood ratio test is developed to test he hypothesis of interest. The asymptotic null distributions of the likeliood ratio test statistic, and of a modified test statistic useful for mproving small-sample properties of the test, are obtained. The resulting heory is extended to the case where 6 psychological tests, each composed f 6 subtests, are compared under a design in which tests i and j , $\leq$ j , have the first i subtests in common, $i,j = 1,2,\ldots,6$. This ork will soon appear as a Research Bulletin and in a chapter in a ymposium volume on Statistical Design and Linear Models.

Testing If Two Measuring Procedures Measure the Same Dimension (772-79)

L. Gleser

PURPOSE: Mr. F. Lord has suggested a convenient statistical technique for testing the hypothesis that two sets of measurements differ only because of errors of measurement and because of differing origins and units of measurement. Our goal is to obtain the exact level of significance of Lord's test, and to study its power function under various alternatives.

PROGRESS AND MAJOR FINDINGS: Lord has indicated that his test is conservative in the sense that the null hypothesis will be rejected incorrectly less often than the reported level of significance value p of the test would indicate. We have shown that in one sense Lord's test is not really conservative since there is a sequence of parameters in H_0 for which the probability of rejection of H_0 under the test tends to p as a limit. However, there is no single parameter value θ_0 in H_0 for which the probability of rejection is exactly p. These assertions are verified by finding lower and upper bounds for the probability of rejection for H_0. Lord has shown that an upper bound for this probability is p, but our upper bound is everywhere less than his bound. Upper and lower bounds are also obtained for the power function of Lord's test, and the exact asymptotic power of this test is obtained (under different definitions of the term "asymptotic"). The above results will appear in a forthcoming Research Bulletin.

Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study (772-82)

D. Rubin
T. Stroud

PURPOSE: To monitor the level of high school grade standards as measured by the regression line of first year university average on final year high school average. Ontario high school graduates wrote province-wide examinations until 1967, but now the schools are on their own. So far the examining standards of high schools have been fairly uniform, and university admissions offices have been successful at using high school grades as the entrance criterion. But will this continue to be the case if the trends toward more flexibility in curriculum bring about wider divergence in examination patterns across schools? To answer this, we are looking at patterns in data from freshman classes at Queen's University.

PROGRESS: Since the least-squares regression line for students from a given school in a given year is unstable due to the smallness of the data set, it is essential to find a satisfactory method of smoothing the parameter estimates for a given school over a sequence of years (and smoothing across schools is necessary as well). Two methods of doing this have been explored and are described under the headings "Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors" and "Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Random Predictors." The random-predictor method has been more successful in that it is considerably less expensive.

MAJOR FINDINGS: In a pilot study (which will be described in a Research Bulletin) involving 8 high schools, 4 matriculation years (1968, 1969, 1970 and 1971) and 565 students, there seems to be evidence that schools' standards were farther apart in 1971 than in the previous three years (although in the years 1968, 1969 and 1970 the schools did not seem any farther apart than in 1967). It is planned that further analyses will be done involving more schools and later matriculation years.

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Random Predictors (772-82)

D. Rubin
T. Stroud

PURPOSE: To provide a usable statistical model for the study described under the heading "Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study." The model is a random-effects two-way MANOVA with interaction with variables X,Y being the final year high school grade average and first year university average, respectively, and with school and matriculation year as variables of classification. This implies regression lines of Y on X with random intercepts varying over the two-way classification, with slopes and residual variances constant. We wish to estimate the parameters for the purpose of seeing how widely schools differ and with a view to forecasting (with standard errors) the regression parameters for a given school for the current year where university grades are not yet available.

PROGRESS AND MAJOR FINDINGS: An empirical Bayes technique has been implemented and run on an APL computer program sequence, using data from eight schools in 1968, 1969 and 1970. The method is relatively inexpensive and seems to yield reasonable results. Although the present program is limited to 24 school-year cells due to memory size in a matrix inversion step, it is a simple matter to change the program to accommodate an unlimited number of cells using an algorithm which has been developed for this purpose. The statistical properties of the empirical Bayes procedure under the assumed bivariate normal model for (X,Y) are yet to be examined.

Three Research Bulletins are being prepared, one describing the method and the data and two others dealing with technical aspects of the algorithms. The method has been described in talks at the University of Massachusetts and at the 1973 meeting of the Statistical Science Association of Canada.

Comparing Regressions When Measurement Error Variances Are Known (772-82) — T. Stroud

PURPOSE: In a multiple (or multivariate) regression model where the predictors are subject to errors of measurement with a known variance-covariance structure, we wish to test two-sample hypotheses of (i) equality of regressions on true scores and (ii) equality of residual variances (or covariance matrices) after regression on true scores.

PROGRESS AND MAJOR FINDINGS: Formulas for calculating the test statistic for each hypothesis are presented, based on Wald's asymptotic chi-square procedure. The formulas are cumbersome, but fortunately one need not use them in practice because the value of the chi-square for given data can be obtained using the computer program AUTEST (Lord, RB-72-42, revised edition; Stocking & Lord, RM-73-7). In a study comparing learning in schools in Portland, Oregon, the chi-square was calculated both directly and using AUTEST, and the results agreed to four significant digits. The research has been written up as RB-73-35.

Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors (772-82)

T. Stroud

PURPOSE: To provide a statistical model for the study described under the heading "Effects of Using High School Grades as the Criterion for College Entrance--A Canadian Study." The model is

$$y_{tsi} = (\alpha_t + u_s + w_{ts}) + (\beta_t + v_s + z_{ts})x_{tsi} + e_{tsi} ,$$

where t is the matriculation year, s the high school, and i the student. y denotes university first-year grade average and x the high school final year average. α and β are fixed-effect parameters, and u, v, w, z are random effect parameters. e represents individual variation. There are unknown variance and covariance parameters as well. We wish to estimate the parameters for the purpose of seeing how widely schools differ and with a view to forecasting (with standard errors) the regression parameters for a given school for the current year where university grades are not yet available.

PROGRESS AND MAJOR FINDINGS: A computer program has been written and is working. It computes the variance-covariance estimates using MINQUE. This is the first stage of the estimation procedure. Several difficulties have emerged: (i) computation is expensive, (ii) the estimates have disconcertingly high sampling variances, (iii) only a few hundred students can be accommodated in one run so the data set must be split and run separately, thus lowering the efficiency of the procedure. The program also requires as input a "norm," or starting solution, which must be obtained by some other method. The Lindley-Novick (e.g., RB-70-32) procedure could provide this.

A description will be written up as a Research Bulletin entitled "Forecasting a Regression Relationship Which Varies over a Large Number of Subpopulations." The model should have applications to areas other than education, e.g., economics. The method has been described in invited talks at Princeton University, Bell Telephone Laboratories and Virginia Polytechnic Institute.

Estimating Variance Components by Combining Unbiased Estimates from Independent Data Sets (772-82)

T. Stroud

PURPOSE: To generalize the practice of taking the positive part of a negative variance-component estimate to the situation where the parameter has been estimated (unbiasedly) independently in n different sources or data sets. We compare the consistent procedure of taking the positive part of the average of the n unbiased estimates with the inconsistent (i.e., asymptotically incorrect) procedure of averaging the n positive parts. The criterion is mean square error divided by the variance of the unbiased estimator, and the unbiased estimator is assumed to be distributed either as a weighted difference of two independent chi-squares or as the large degrees-of-freedom limit of this, i.e., Gaussian.

PROGRESS AND MAJOR FINDINGS: When the number of sources $n = 2$, 3, or 4 and when the positive chi-square has 20 or more degrees of freedom, the inconsistent procedure is substantially better, provided the coefficient of variation of the unbiased estimator is less than two. An example where this occurs is the one-way components of variance with, e.g., 26 groups per source and 10 observations per group and a between-group variance of at least 1/35 of the within-group variance.

This research was motivated by the situation in "Estimating and Forecasting Regression Parameters in an Achievement Prediction Model--Fixed Predictors" where the splitting of the data into several computer runs requires the combining of estimates of variance components, some of which may be negative. This research was presented at the 1973 Eastern Regional Meeting of the Institute of Mathematical Statistics, and is available as RB-73-47.

Use of INDSCAL in Assessing a Cerebral Deterioration Pattern in Psychiatric Patients (772-82)

T. Stroud

PURPOSE: To see if the widely-used INDSCAL algorithm (Individual Differences Scaling--Carroll and Chang) sheds any light on dimension or factor patterns in the Wechsler Adult Intelligence Scale (WAIS) with respect to a population of psychiatric hospital patients.

PROGRESS: The INDSCAL computer program was run using diagnostic groups in place of individuals, and measuring dissimilarity between subtests by sums of squared differences of standard scores over all patients in the diagnostic group.

MAJOR FINDINGS: In the three-dimensional solution, the first dimension has been identified as verbal vs. spatial/quantitative, the second as memory-ingenuity, and the third as "hold-don't hold." The labelling of the third dimension refers to the use of WAIS subtest scores to indicate cerebral deterioration, and is based on contrasting those subtests which deteriorate with age with those which do not. The third dimension loadings showed a correlation of .885 with the subtest weights in a deterioration coefficient due to J. S. Lawson which seems to perform better as a deterioration indicator than does Wechsler's Deterioration Index.

The diagnostic groups that scored as most heavily weighted on the "hold-don't hold" dimension were the senile group and alcoholic psychosis. Alcoholism and drug dependency, on the other hand, were among the least heavily weighted, along with personality disorder, mental retardation, epilepsy and neurosis.

Standard Errors for Rotated Factor Loadings (772-83) — R. Jennrich

PURPOSE: In spite of the wide scale use of factor analysis by psychometricians and statisticians in general, and the widespread availability of factor analysis programs, the assessment of standard errors of measurement for the estimates these programs produce has received very little attention. Literally thousands of factor loadings are computed every day, but not a single standard error, at least not in the predominantly most popular case of analytically rotated loadings. An attempt to remove this defect in standard factor analysis programs is being made by developing formulas for the asymptotic standard errors of factor loading estimates applicable to the methods of estimation and rotation most commonly used.

PROGRESS: Results have been summarized in five Research Bulletins. The first (RB-73-17) derives asymptotic standard errors for orthogonally rotated factor loading estimates. These apply to factor analysis or principal components analysis and to any of the complete class of invariant fourth-degree orthogonal rotation criteria. The second (RB-73-28) obtains similar results for the complete class of invariant fourth-degree oblique rotation criteria. The third (RB-73-31) reports an error in Lawley's formulas for unrotated loadings upon which the first two Research Bulletins were originally based. This RB corrects Lawley's derivation and formulas. In the fourth (RB-73-40), simplified standard error formulas are obtained by inverting an augmented information matrix. These are computationally more expensive and apply, easily at best, only to maximum likelihood estimation with orthogonal rotation. Their use, however, uncovered the error reported in the previous RB. The last (RB-73-43) gives a nontechnical summary of the statistical behavior of factor loading estimates and identifies singularities which lead to pathologies. It explains, for example, why surprisingly stable rotated loadings may arise from unstable initial loadings. At the present time the first two RBs have been accepted by Psychometrika, and the last by the British Journal of Mathematical and Statistical Psychology.

MAJOR FINDINGS: Methods for computing standard errors of estimate for analytically rotated factor loading esimates on a routine basis are now available. In applications these standard errors seem to be pleasantly, almost surprisingly, small; often ranging from .05 to .10 for rotated loadings computed from a correlation matrix. In this case, the reciprocal of the square root of the sample size seems to give a crude but useful approximation to the standard error of estimate.

Developmental Study of Attention within the First Two Years of Life (NSF, 776-3)

M. Lewis
E. Scott
C. Wilson

PURPOSE: This study was designed to study response decrement and recovery in the first two years of life as affected by a number of factors. Repeated and novel stimuli were presented to each subject in two modalities--the auditory and the visual. Each subject received a simple series and complex series in both modalities. The type of change produced by the violation stimulus was changed. Thus, we are able to compare response decrement and recovery across age, across modalities, across complexities, or across stimulus violations. In addition, measures of cognitive development (the object permanence scale of Escalona and Corman and the Bayley Mental Development Scale) were obtained.

PROGRESS: Both a cross-sectional and a longitudinal sample of infants were seen at 3, 6, 9, 12, 18, and 24 months of age. Each subject made two visits--on one the subjects received a visual complex and auditory simple series; on the other, an auditory complex and visual simple (the order was counterbalanced). Measures of fixation, cardiac response, and activity were recorded for the visual episodes. Only cardiac and activity data are available for the auditory episodes. Data analysis is not yet complete for all ages.

MAJOR FINDINGS: A detailed cross-modality comparison of the data collected at three months was carried out. Complexity and sex differences were also considered. The data indicate several trends within and across modalities. These are summarized as follows. For visual stimuli (measured by fixation time), boys showed discrimination on the basis of complexity while girls did not. For auditory stimuli (measured by cardiac deceleration), neither boys nor girls showed discrimination although there is a tendency towards discrimination in the girls' data. All of the discriminations are in the direction of more attention to complex than to simple events. Total amount of cardiac deceleration indicates a sex x modality difference. Girls' deceleration to auditory stimuli is greater than the decelerations associated with any other events, including boys' and girls' visual and boys' auditory. Response decrement across the six trials was found to be minimal in all areas of measurement. There appear to be two important trends in the data: (1) greater stimulus discrimination in the visual than the auditory mode, and (2) greater cardiac response for girls to auditory than to visual events. However, difficulties arise in comparing the nature of the stimuli in the two modalities. It may be that the simple and complex auditory stimuli are not as distinct on the complexity continuum as the visual stimuli. Thus the greater discrimination found in the visual mode may be a function of inherent stimulus characteristics rather than modality differences. These considerations point to the problems that arise in comparisons across modalities. The finding that girls show greater deceleration in the auditory mode than in the visual mode does not suffer from such theoretical difficulties. Our conclusion is in accord with earlier work (Kagan & Lewis, 1965; Watson, 1969) which suggests that such

stimulus-modality differences do exist, with girls responding more to auditory than to visual events. Thus comparisons across modalities are informative, but complex and difficult to interpret. Much more empirical evidence is clearly needed before conclusive inferences can be drawn. A paper, presenting the findings to date, was presented at the Eastern Psychological Association meetings in New York in April 1971.

Further analyses were concerned with the questions of the pattern of change of response decrement and recovery across all the ages studied. In addition the questions of individual consistency across age and within and across modalities were explored. Response decrement of the fixation measures to the visual stimuli showed an age effect, with older infants showing more response decrement than younger infants. This result confirms our earlier findings which indicated that amount of response decrement was age related. The second result has to do with stimulus complexity. At each age level except six months there is a complexity effect with the more complex stimuli eliciting less response decrement than the simple stimuli. Again this was not unexpected since stimulus complexity has been repeatedly shown to affect attentional distribution. The decrement in HR deceleration also shows a stimulus complexity effect. Except for 24 months the simple stimulus elicits more response decrement than the complex one.

In general, then, age related response decrement has been demonstrated with two different visual stimuli varying at least in complexity. This age related response decrement appears invariant with respect to the nature of the stimulus. That is, stimulus complexity effects were to be found in the amount of response decrement but not in the age relationship. When the longitudinal data are examined for evidence of consistency across age, the fixation data for the visual data are clearest and indicate a consistency in amount of response decrement across the first 18 months of life. The HR data, both for the visual and auditory episodes, in general, fail to indicate any consistency across age. Given the results of the fixation data we can only conclude that HR during an attentional task is not consistent across age. Whether this is due to the large variability usually encountered in such studies, nature of measurement, or too small sample size remains to be determined.

The longitudinal data were also analyzed to discover if response decrement was consistent within modalities across ages, and also to see if a relationship existed between modalities. The data fail to support within visual modality consistency; however, there is support, albeit limited, for within auditory modality consistency. Across modality consistency in response decrement was strongest for the two episodes which elicited the most regular age effects--namely, visual complex-auditory simple series.

The pattern of response recovery across ages differed with the type of violation introduced. The changes were from straight to curved (curvature); color to black and white (color); and a change in the number of lines. There were three developmental patterns: a linear increase with

some asymptotic leveling for curvature, an exponential function for color and a curvilinear function for number violation. This same pattern is visible when looking at percentage of subjects showing fixation recovery.

The HR recovery data by specific violation show little of the pattern that we saw for the fixation data. The only consistent pattern is that for the 3 line curvature change, which shows the greatest HR recovery as it did for the fixation recovery. Once again the cardiac response proves to be only marginally related to the fixation data. The auditory recovery data are comprised only of HR deceleration. Overall the most effective violation is the change in tone from C to F. This change produced the most response recovery while the least effective was the octave change.

Over all ages there was consistency in fixation recovery. Elimination of the 3 and 6 month data when little recovery occurred results in an even greater degree of concordance. Similar results were found for heart rate recovery consistency, especially when the 3, 6, and 9 month data were not considered.

For the auditory episodes the heart rate recovery data failed to show any consistency across age even when the earlier ages were not considered. Tests for within and across modality consistency showed simply that response recovery to one type of violation was not necessarily related to recovery to other types of change.

These findings were presented at the XX International Congress of Psychology in Tokyo, Japan, August 1972.

M. Lewis
H. McGurk

Evaluation of Infant Intelligence: Infant Intelligence Scores--True or False? (NSF, 776-3; Spencer, 815-1)

PURPOSE: The purpose of this study was to investigate the predictive value of various tests of infant cognitive development, and the relationships among the various tests. The suitability of administering general tests to evaluate specific infant intervention programs was also considered.

MAJOR FINDINGS: Twenty infants were tested at 3, 6, 9, 12, 18 and 24 months. The tasks administered included the mental scale of the Bayley Scales of Infant Development and the object permanence scale from Escalona and Corman's Scales of Sensory-Motor Development. At 24 months, language production and comprehension tasks based on items selected from the Peabody Picture Vocabulary Test were also given. Results indicate that there is no reliable relation between measures of infant intelligence during the first 24 months of life. Relationships among the tests seem to exist only when similar types of items appear on the different tests. No evidence was found to support the view that intelligence is a single unitary capacity which unfolds at a steady rate throughout the process of development. The data also cast doubt on the notion that scores on infant intelligence scales can be generalized beyond the particular set of abilities sampled by the items administered at the time of testing. The implications of these conclusions for evaluation of infant intervention programs seem clear. Simply stated, infant intelligence scales are unsuitable instruments for assessing the effects of specific intervention procedures. This is true primarily because infant intelligence is not a general, unitary trait, but is, rather, a composite of skills and abilities that are not necessarily covariant.

Developmental Model of Cognitive Style (779-3) A. H. Baker

Study 1: Relations between cognitive controls and school performance in the early grades.

PURPOSE: To identify principles by which children organize and process incoming information from the environment and to relate these individual differences in perceptual-cognitive style to academic success in the early school years. This research is being framed within the context of a particular developmental model, according to which all cognitive controls are not equally available at all stages of life, and a newly formed cognitive control does not replace one that was formed earlier. Rather, the model views cognitive controls as forming a hierarchy of organizations of cognitive behavior, and makes use of the concept of levels within each organization. Focal Attention (i.e., directing attention actively to stimuli and scanning broad sections of the environment so that unique properties of objects are registered) is assumed to emerge first in development. The principle of Field Articulation (i.e., attention directed selectively to relevant stimuli and withheld from irrelevant stimuli guided by the central requirements of the task) emerges next, subordinating and integrating Focal Attention as a distinct process. Leveling-Sharpening (i.e., maintaining in memory the organization of relevant-irrelevant information in the form of a differentiated image with which subsequent information is related and compared) is the third cognitive control principle to emerge and is assumed to subordinate and integrate the preceding two principles. Finally, Equivalence Range is the fourth principle whereby symbols or labels are introduced in order to interrelate bits of information allowing objects to be dealt with in terms of categories or classes. In brief, the present research is aimed at exploring how each of the above described cognitive style dimensions relates to school performance in kindergarten and first grade children.

PROGRESS: One hundred fifty white, low to middle class kindergarten children have been given a battery of procedures which assess the above-described cognitive controls, prereading and reading skills, ratings of general classroom behavior (by the teachers), impulse control, and selected personality measures.

It is planned to relate a child's cognitive developmental status, in terms of information processing strategies, with the status of his cognitive skills in terms of reading. The goal is to search for patterns of information processing strategies (and their level of development) associated with the emergence of effective reading skill and to learn more about the relation of certain aspects of personality here presumed to be critical in the learning process and the early formation and operation of cognitive styles.

Fifty first grade children who had been evaluated one year ago were also recently given the same battery of procedures outlined above. The focus here was upon (a) the stability or change in patterns of the cognitive structures and skills under study, and (b) the interrelationship between the assessments of academic success obtained from the teacher and the cognitive status observed during the present testing as well as that observed a year ago.

Data collection and processing phases have been completed for both aspects of this study and data analyses are now in progress. (Study done in collaboration with Sebastiano Santostefano, Boston University School of Medicine.)

Study 2: Individual differences in attentional processes in adults and children.

PURPOSE: Recent research has examined the hypothesis that individuals can be ordered along a dimension in terms of their tendencies to reduce or to augment the intensity of persistent incoming stimulation. At one extreme, the reducer attenuates the intensity of incoming stimulation, thus enabling him to handle high intensities of stimulation quite well but making him acutely uncomfortable when environmental stimulation is minimal (e.g., under sensory deprivation). At the other extreme, the augmenter magnifies the intensity of incoming stimulation, thus showing intolerance of high levels of stimulation but high capacity to cope with very low stimulus intensity. In between is the moderate, who neither greatly enlarges nor minimizes incoming stimulus intensity. Measurement of this dimension has involved variants of the Kinesthetic Figural Aftereffects Task (KFA). Individuals are designated as reducers, as moderates, or as augmenters if a standard block is judged respectively as smaller, as about the same size, or as larger after intervening stimulation. Findings have indicated, for example, that reducers are more tolerant of pain, less able to endure sensory deprivation, more likely to show reduced averaged cortical and visual evoked responses to strong light stimulation, and more likely--if male--to participate in "contact" athletics such as football.

In a series of interrelated studies, the present research seeks to explore the following substantive-theoretical and methodological issues suggested by prior KFA research: (1) Most differential studies utilizing KFA procedure can be criticized either on the basis of the psychophysical procedures employed and/or on psychometric grounds. The first phase of the present research will therefore simultaneously reexamine the reliability and validity of two major variants of the KFA procedure, using normal adults as subjects, and if successful, (2) an attempt will be made to delineate more precisely the nature of the psychological processes underlying individual differences in performance on KFA tasks. (3) A third phase will attempt to delineate educational implications of the augmentation-reduction dimension. The augmentation-reduction typology suggests that the optimum learning environment may be quite different for these three kinds of individuals. The reducer should learn best in an environment in which the techniques that introduce educational materials utilize fairly high levels of sensory stimulation, in which a variety of educational approaches are utilized and in which stimulation is provided by frequent social interaction. However, a calm, nonstimulating environment should best promote learning by the augmenter, who needs far less stimulation than does the reducer and, in fact, would be quite uncomfortable in highly stimulating surroundings. Augmenters should also profit from educational projects which they are given to carry out on their own. (4) Finally, this research will focus upon age changes in the hypothesized augmentation-reduction dimension, and in the correlates of this dimension, with special reference to other possible dimensions which may be tapped by the KFA task.

PROGRESS: Since its inception in the summer of 1970, the following has been achieved: (1) an extensive literature search has been accomplished; (2) a number of substantive and methodological issues which appear to be critical in studying individual differences in kinesthetic figural aftereffects have been tentatively identified, and a research program aimed at exploring them has been broadly sketched; (3) equipment has been secured for administering the two principal variants of the KFA task; (4) an extensive battery of tests and tasks aimed at exploring the issues described above has been assembled and/or developed; (5) personnel have been recruited and trained in the administration of the two types of KFA tasks and of the test batteries; (6) ten different studies, aimed at exploring the goals of the present research delineated earlier, have been completed.

The findings to date have both clarified a number of substantive and methodological issues, while at the same time expanding our conceptualization regarding this research area. Substantively, it is now clear that individual differences in the KFA task reflects both trait variance and state variance. Thus, as a personality dimension, KFA scores predict individual differences in thinking about the future, time perspective, and attitudes toward death, findings which hold not only for an undergraduate population but also for a community-active geriatric group. At the same time, however, a number of aspects of organismic state--including the degree of self-rated tiredness, day in the monthly menstrual cycle, intake of various medications, presence or absence of illness or pain etc.--affect scores on the KFA task. Methodologically, it has been shown that when such "state" variables are controlled, there is a significantly greater positive test-retest reliability than when state is not controlled. Moreover, certain problems in the use of a test-retest design in the study of kinesthetic figural aftereffects have been delineated.

In expanding our interests regarding the ways in which people react to the intensity of incoming environmental stimulation, we have extended our attention beyond the Kinesthetic Figural Aftereffects task. As a result, our efforts during the past year have focused on devising a very extensive battery of tasks which would assess this dimension in multiple modalities using multiple methodologies, while at the same time including in this battery a number of the more traditional cognitive style variables to clarify possible interrelationships between the present dimension and the dimensions studied by other researchers. As a result of this effort, we have devised a highly diversified but very specialized 12-hour battery of tests which includes not only extensive paper-and-pencil measures but a wide variety of psychophysical tasks pertinent to the present problem area. This battery has been administered to 120 college-age subjects; all data has been processed, and data analyses are in progress.

Another line of investigation has focused upon individual differences in the aging process. Because the reducer tends to damp down incoming stimulation, it has been hypothesized and found in carlier studies that he shows greater tendency to seek more intense and more complex levels of stimulation. We therefore hypothesized that (1) reducers would show a higher level of social

involvement and (2) the reducer would maintain his higher level of social interaction during old age whereas the augmenter would show more tendency to withdraw from social interaction. Both predictions have been confirmed in a recent self-report type study.

Results of some of the above findings have been or will be reported as follows: (1) Individual differences in the psychology of death. To be presented at meetings of the American Psychological Association, August 1973, Montreal, Quebec, Canada. (2) Effects of variation in ongoing and induced organismic state upon kinesthetic figural aftereffects. Presented at meetings of the Eastern Psychological Association, May 1973, Washington, D. C. (with Brian Mishara, Irene W. Kostin, and Laurence Parker). (3) Will to live, future planning, and kinesthetic figural aftereffects: Towards a theory of individual differences in the aged. Presented at the 25th Annual Meeting of the Gerontological Society, December 1972, San Juan, Puerto Rico (with Brian Mishara and Irene W. Kostin). (4) Do people who seek less environmental stimulation avoid thinking about the future and their death? A study of individual differences in kinesthetic figural aftereffects. Presented at meetings of the American Psychological Association, September 1972, Hawaii (with Brian Mishara and Irene W. Kostin). (Study done in collaboration with Dr. Brian Mishara of Socio-Technical Systems Associates, Boston.)

Attachment Behavior in Thirteen-Month-Old Opposite Sex Twins (779-4; Spencer, 815-1)

J. Brooks
M. Lewis

PURPOSE: To study whether two children, a male and a female, show differential degrees of attachment behavior toward the same mother.

PROGRESS AND MAJOR FINDINGS: Seventeen sets of opposite sex twins, 13 to 14 months old, were observed in a playroom situation with their mothers. Attachment behaviors, toy preference, style of play, and activity level were recorded. Analysis of four attachment behaviors indicated that girls looked at, vocalized to, and maintained proximity with their mothers significantly more often than did their brothers. Girls also touched their mothers more often, although this difference was not significant. The style of play and activity level measures revealed no sex differences. In terms of specific toy preferences, girls tended to play with the cat and dog more, while boys tended to play more with the pull toy. No other differences in toy preference were found.

The effects of social class, activity level, and separation from the mother on infant attachment behavior are discussed. The importance of the mother as a socializing and sex-typing agent is stressed.

A paper was presented at the Society for Research in Child Development meetings, Philadelphia, March 1973, and is available as RB-73-25.

The Integration of Faces and Voices in Infancy
(779-4)

M. Lewis
L. Rosenwein
C. Wilson
H. McGurk

PURPOSE: Two studies were done under this rubric. The general purpose involved the question of whether intermodality integration exists in early infancy. Most theories of perceptual development postulate a gradual intermodality integration with age. We were intrigued by Aronson and Rosenbloom's (Science, 1971) research showing that one-month-old babies were disturbed by a mismatch between their mothers' faces and the displacement of their voices. The first study was designed to replicate this finding in a more controlled experiment. The second study was designed to elaborate and extend theoretical issues investigated in the first study. We wanted to know whether learned information can be coordinated between two senses. In order to determine the answer to this question we explored the infant's reaction to a stranger's voice mismatched with his mother's and vice versa as well as to face and voice mismatches involving location.

PROGRESS: Data have been collected and analyzed in both studies. A preliminary report is available and is in the process of being submitted for a Research Bulletin.

MAJOR FINDINGS: Neither study 1 nor 2 replicated Aronson and Rosenbloom's results. On the other hand, one-month-old infants as well as four-month-old infants were able to distinguish between normal face-voice integrations and discrepant ones of all types (learned as well as unlearned). Additionally, we found evidence of increasing sound localization with age.

Infants' Responses to Pictures of Mother, Self, and Other (779-4)

M. Lewis
J. Brooks

PURPOSE: This study was designed to explore further the infants' ability to discriminate between familiar and unfamiliar persons and to differentiate among strangers. We recently found that infants exhibit negative affect (fear and wariness) to adult strangers and positive affect to a child stranger (Lewis & Ban, 1972). In addition, infants responded quite positively to their mirror image as well as to their mothers. The present study extends our inquiry into the infant's ability to recognize himself as well as differentiate among various strangers when presented two-dimensional representations.

PROGRESS AND MAJOR FINDINGS: Pictures of the infant's mother, the infant himself, other infants, 5- and 10-year-old children, and adults were shown to 54 infants in four age groups: 10, 12, 16, and 18 months. Pictures of both male and female strangers were presented for the baby, children, and adult stranger conditions. In general, the older infants were more likely to differentiate among the pictures than were the younger infants. Differentiation was defined in terms of fixation, affect, and spontaneous labels. The concept of self is discussed in terms of the data. A sex of subject-sex of stimulus interaction was also found which was discussed in terms of gender identity. A chapter based on this study will appear in L. Cohen and P. Salapatel (Eds.), Perception in Infancy.

Longitudinal Study of Infant Attention and Its Relationship to Subsequent Perceptual-Cognitive Development (779-4)

M. Lewis
C. Wilson

PURPOSE: This longitudinal study investigates attentional processes during infancy and their relationship to cognitive growth in the preschool period. Specifically, the study explores: (1) problems in measuring attention; (2) factors influencing attention, such as sex, and clinical problems (birth trauma); (3) stimulus differences such as facial vs. nonfacial and familiar vs. novel; (4) developmental consequences of differential attention, including the relationship of attentional differences in infancy to preschool performance in concept formation tasks, learning problems, reading, and language acquisition.

PROGRESS: To explore these problems, two samples of 32 infants, 16 boys and 16 girls, were seen at 6, 13, 25 and 44 months of age. At each age level, a variety of visual and auditory stimuli were presented to the infants and a large number of response measures were taken. In addition to their response to visual and auditory stimulation, a wide variety of other behaviors were observed. These included such behaviors as concept formation, problem solving, task persistence, language acquisition, play behavior, and IQ measures.

The data, yet to be fully analyzed, will be used to explore the major areas of interest outlined below.

(a) Measurement of attention. The responses observed have been: (1) fixation (receptor orientation), (2) vocalizing, (3) crying, (4) autonomic reactivity (specifically, heart rate and respiration) and (5) gross motor activity. Investigators often harbor implicit assumptions about each of these response indices. For example, differential fixation time is often assumed to reflect a capacity for differentiation or discrimination. More significantly, absence of differences in fixation time is believed to reflect the infant's inability to differentiate among the visual patterns being presented. This inference may be fallacious, for often the infant fixates for an equal period upon two stimuli but will only vocalize or smile to one of them. Vocalizing has been assumed to indicate both a specific social response and general arousal level. Similarly, gross motor activity or heart rate acceleration is often regarded as an indicator of activity or arousal level. It is clear that detailed empirical work is necessary to assess the viability of these assumptions. At present, it appears that each of the infant's responses viewed singly is probably ambiguous in meaning and can reflect different psychological processes. A long fixation time can occur to a familiar stimulus or to an uncertain one. Cardiac deceleration can occur when the infant studies the stimulus and is quiet, but acceleration often appears when he studies the stimulus and is active. Thus the present study intends to contribute information relevant to understanding the varied responses associated with distribution of attention to auditory or visual patterns. Analysis of six-month data indicates that attention to visual and auditory stimuli is usually associated with the following type of response pattern: receptor orientation; decrease in gross motor activity, crying, and vocalizing; autonomic responses of cardiac deceleration; and a slowing of

respiration. Individual differences in this response pattern have been observed. Developmental differences in this response pattern have yet to be explored.

(b) Individual differences in attention. Marked individual differences have been found in the attentional behavior of six-month-old infants. Are these differences consistent in the visual and auditory modalities, and are the differences stable over age?

(c) Stimulus differences in attention. Because a variety of visual and auditory stimuli were presented at each of the four age levels, it will be possible to explore the influences of different stimuli on attentional behaviors. Both visual and auditory stimuli can be divided into social and nonsocial, i.e., human forms, faces, or voices versus nonhuman patterns or sounds. These two classes of stimuli were presented at all age levels and the study seeks to explore the effects of these stimulus qualities on the attentional process, i.e., their ability to attract (length of time oriented toward) the infant's attention, and the ability to elicit social responses, i.e., vocalization and smiling.

(d) Relationship between attentional measures and cognitive capacities and styles. Along with the various attentional measures to visual and auditory stimuli, a variety of other cognitive capacities and styles were observed for the one-, two-, and three-year levels. These include: (1) language tests at two and three years, (2) play behavior at one, two and three years, (3) concept formation at two and three years, (4) problem solving at two and three years, and (5) IQ scores at three years. The cognitive style measures of impulsivity and reflectivity as measured by MFF task and persistence task were given at three years. It is therefore possible at each age level to investigate the relations among these various measures of cognitive capacity and to relate them to the attentional measures. For example, there is evidence that amount of fixation is related to play behavior at one year and that IQ scores can be predicted from attentional scores at the three-year level.

(e) Personality development. In the longitudinal investigation of attention and cognitive growth, several personality characteristics were assessed. Observation of individual differences in attentional behavior revealed that personality characteristics might covary with them. Personality measures obtained at the one-, two- and three-year levels were: (1) measurement of the infant's vigor, (2) dependence behavior as measured by interaction of infant with mother or teacher, (3) sex role appropriate behavior measured in play behavior, and (4) task persistence. Preliminary observations suggest that a dependent child may spend less time looking at nonsocial stimuli as compared to social stimuli. Moreover, when left alone in a relatively unfamiliar room, his attention span might be more a function of wanting to see the mother than of interest in the stimulus being presented. It was also observed that vigorous children had more difficulty attending than placid, less vigorous ones who showed little physical activity. Analysis of the data is proceeding along all domains of interest.

MAJOR FINDINGS: (a) Measurement of attention. Principal component analyses on the attention data indicate that the configuration of two components, together accounting for more than half the total variance, is quite consistent both for the different types of attentional stimuli and across age. The first component has been called an orienting component of attention, as it represents those responses involved in orienting to the stimulation: fixation, cardiac deceleration, and a reduction in gross activity. The second component is interpreted as an affect component, representing the vocalization and smiling responses to the attentional stimuli.

(b) Individual differences in attention. In general the results show little consistency of individual differences in attention, either between the visual and auditory modalities, or across the four age points. No consistent sex differences were found. It appears that the method of assimilating environmental information varies as a function of the type of information to be assimilated. Moreover, patterns of attention vary with age, perhaps in the search for the most efficient method of processing incoming information.

(c) Stimulus differences in attention. The analysis of the data for the social and nonsocial stimuli separately revealed few consistent differences. However, principal component analyses on the responses to the two kinds of stimuli indicated a stronger (accounting for a greater percent of the variance) orienting component for the nonsocial than for the social stimuli, and a stronger affect component for the social than for the nonsocial stimuli. Thus, the difference in the response to the two types of stimuli is not so much a difference in quantity as in the relative predominance of certain response measures such as vocalizing and smiling.

(d) Relationships between attentional measures and cognitive capacities and styles. There appears to be no consistent relationship between the attentional and cognitive measures, indicating that the processing of incoming information (attention) is separate from cognitive capacity. In view of the lack of consistency in patterns of attention it is not surprising that the attentional and cognitive measures are not related. Even if cognitive capacity is assumed to be constant (a tenuous assumption at best), the variations in attention as a function of both type of stimulus and age preclude the finding of consistent relationships between attention and cognition.

Mothers and Fathers, Girls and Boys: Attachment Behavior in the First Two Years of Life (779-4)

M. Lewis
M. Weinraub
P. Ban

PURPOSE: In this longitudinal study of one- and two-year-old children, the child's expression of four attachment behaviors--touching, looking at, vocalizing to, and proximity to each parent--was examined. We were interested in observing (a) how these four behaviors are interrelated both within and across ages, (b) sex of child and sex of parent differences in the use of attachment behaviors, and (c) the relationship between the patterns of attachment behavior expressed to each parent and measures of the child's intelligence.

PROGRESS AND MAJOR FINDINGS: A small group of children was observed individually at one and two years of age in a naturalistic free play situation in the presence of their mothers and again in the presence of their fathers. Intelligence tests were administered to the children.

There were no overall child sex differences in the expression of attachment behaviors toward the parents at either age. There were parent differences at age one, with more attachment behavior directed toward the mother than toward the father. By two years of age these parent differences in the expression of sex differences disappeared for the most part.

The data suggest that, in general, there may be a transformation from the use of proximal (touching and proximity) attachment behaviors to the use of distal behaviors (looking and vocalization). However, both the sex of the parent and sex of the child influences the clustering of attachment behaviors. Measures of infant intelligence were found to be correlated with the changing patterns of attachment. More intelligent children tended to show greater decreases in their expression of proximal attachment behaviors and greater increases in their expression of distal attachment behaviors than less intelligent children.

The sex of child and sex of parent differences we observed in the expression of attachment behaviors by two-year-old children seem to be precursors of sex differences in interpersonal behaviors characteristic of older children and adults. A report of this study was presented at the meetings of the Society for Research in Child Development in Philadelphia, April 1973.

Differentiation and Development of Social Norms (779-7) — W. Emmerich

PURPOSE: The study examines developmental trends in age-role and sex-role normative differentiations and relates these trends to developmental theories of social learning, cognitive complexity, and cognitive structure.

PROGRESS: Shared beliefs on how the self should act toward others (norms) were studied cross-sectionally in 680 middle-class children and adolescents. Subjects judged the normative import of statements differing in (1) behavioral content (agree with other, help other, seek help from other, argue with other); (2) sex and generation of reference figures holding norms for the self including the self (sources), and (3) sex and generation of recipients of the self's actions (objects).

MAJOR FINDINGS: The four behavioral contents were differentiated similarly at all ages (continuity), but age-role and sex-role normative differentiations of sources and objects exhibited marked developmental trends (discontinuities), often generalizing across behavioral contents. During middle childhood higher standards were expected in relationships with parents than with peers irrespective of source, whereas during adolescence this differentiation was reversed for attributions to sources within the child's generation. Sex-role norms were strongly sex typed during middle childhood, a pattern that decreased during later childhood and adolescence. Related developmental shifts occurred in choices of identification models, indexed by convergences between norms held by the self and those attributed to other sources. Certain developmental trends were accelerated in brighter subgroups. The study is published in the Journal of Personality and Social Psychology (with K. S. Goldman and R. E. Shore), 1971, 18, 323-353.

The study provided information on developmental trends in extreme response choice, found to decrease linearly with age in both sexes and earlier (curvilinearly) in girls. Cognitive level was found to mediate choice of an absolute and extreme category, but not choice of a non-absolute extreme scale category. This aspect of the study was published in Developmental Psychology, 1971, 5, 540.

In a follow-up study, middle-class children and adolescents evaluated single trait descriptions of hypothetical persons of their own sex. Developmental trends in trait evaluations were examined and related to the subject's sex, cognitive ability, and generational status of the hypothetical stimulus person (adult vs. peer). Certain findings were most consistent with ego-developmental theory, while others were more consistent with a person-perception analysis or with a social desirability interpretation. Popularity and similarity judgments were partially differentiated from other components of value such as like-dislike. Evidence suggested that "personal" evaluations become increasingly differentiated from "social" desirability judgments with development, and that self-esteem is maximally

Variance and Invariance in the Mother-Infant Interaction: A Cross-Cultural Study (779-4)

M. Lewis
P. Ban

PURPOSE: To compare mother-infant interaction across different cultures.

PROGRESS AND MAJOR FINDINGS: Samples of Yugoslavian and American mother-infant dyads were observed in detail. In addition, data collected on Dutch, Zambian, and Senegalese mothers and their three-month-old infants made comparisons possible across five different cultures. In general there was considerable agreement in terms of the caregiving the infants received. Consistency across maternal behavior was first discussed in terms of its relationship to the invariances that exist across human cultures. The fact that behavior studied out of context may result in serious misinterpretation was also discussed.

A paper was presented at the Burg Wartenstein symposium, Cultural and Social Influence in Infancy and Childhood, Burg Wartenstein, Austria, June 1973, and is available as RB-73-27. It will also appear in P. H. Leiderman and S. R. Tulkin, Cultural and Social Influences in Infancy and Early Childhood. Stanford: Stanford University Press, in press.

engaged when the individual evaluates himself on traits defining stage-related themes in ego development. This study will be published in Child Development.

Findings from the above studies related to sex-role development are incorporated in a chapter entitled "Socialization and Sex-Role Development" to appear in P. B. Baltes and K. W. Schaie (Eds.), Life-Span Developmental Psychology: Personality and Socialization. New York: Academic Press, 1973.

On Proportional Profiles in Factor Analysis
(NICHHD, 780-8; 772-63)

W. Kristof

PURPOSE: It is intended to examine a number of problems connected with proportional factor profiles. Existence questions, ways of transformation and goodness of fit criteria are studied. Such problems arise, e.g., in multivariate selection situations.

PROGRESS: A number of pertinent theoretical results have been obtained. Their incorporation in a more complete theory is sought.

MAJOR FINDINGS: Necessary and sufficient conditions for the existence of proportional factor profiles on the basis of factor matrices and factor patterns are established. The merits of different goodness of fit criteria are evaluated. The problem of proportional patterns is reduced to the problem of proportional profiles involving factor matrices. The methods employed are based mainly on eigenvectors and eigenvalues.

On a Statistic Arising in Testing Correlation (NICHHD, 780-8)

W. Kristof

PURPOSE: The paper is devoted to the study of a certain statistic, u, defined on samples from a bivariate population with variances σ_{11}, σ_{22} and correlation ρ. If the corresponding statistics are s_{11}, s_{22} and r, then $u = 2r\,(s_{11}s_{22})^{\frac{1}{2}}\,(s_{11} + s_{22})^{-1}$. The possible relevance of this statistic in mental test theory is to be explored.

PROGRESS: The work is contained in RB-71-21. It has been published in Psychometrika, 1972, 37, 377-384.

MAJOR FINDINGS: Let the parameter corresponding to u be υ. Under binormality conditions the following is shown. (i) If $\sigma_{11} = \sigma_{22}$, then the distribution of u can be obtained rapidly from the F distribution. Statistical inferences about $\rho = \upsilon$ may be based on F. (ii) In the general case, allowing for $\sigma_{11} \neq \sigma_{22}$, a certain quantity involving u, r and υ follows a t distribution. Statistical inferences about υ may be based on t. (iii) In the general case a quantity t' may be constructed which involves only the statistic u and only the parameter υ. If treated like a t distributed magnitude, t' admits conservative statistical inferences. (iv) The F distributed quantity mentioned in (i) is equivalent to a certain t distributed quantity as follows from an appropriate transformation of the variable. (v) Three test statistics are given, which can be utilized in making statistical inferences about $\rho = \upsilon$ in the case $\sigma_{11} = \sigma_{22}$. A comparison of expected lengths of confidence intervals for ρ obtained from the three test statistics is made. (vi) The use of the formulas derived is illustrated by means of an application to coefficient alpha.

Cognitive Styles and Psychopathology (NICHHD, 782-3) S. Messick

PURPOSE: Recent studies of perceptual and cognitive response patterns of schizophrenics have indicated striking differences between groups of patients with different symptom patterns (paranoid vs. nonparanoid), different levels of adequacy in life adjustment prior to illness (good premorbid vs. poor premorbid adjustment), and different lengths of institutionalization (acute vs. chronic). The present study attempts to utilize the extreme forms of cognitive organization and various personality trait patterns.

PROGRESS: An extensive battery of perceptual, cognitive, and personality measures has been administered to a sample of 95 hospitalized schizophrenics. These procedures were administered either individually or in small groups of four or five patients. The battery included measures of perceptual field-independence (Rod-and-Frame Test, Embedded-Figures Test), element articulation, form articulation, scanning, verbal closure, perceptual closure, extremity and confidence of judgment, category width, risk taking, leveling-sharpening, conceptual differentiation (object sorting), ideational fluency, and verbal comprehension. Also included were items of case history information; a personality inventory assessing several dimensions of neuroses and character disorder; and several rating scales for symptom expression, properties of delusions, and suddenness of onset of illness.

The sample contained roughly equal numbers of paranoids and nonparanoids, who in turn were approximately equally divided into acute and chronic subgroups (using 18 months of hospitalization as the dividing point). The scoring of these measures has been completed, and their reliability for this sample has been appraised and found satisfactory. Various analyses are under way to evaluate subgroup differences on cognitive style and personality measures (in a 2 x 2 breakdown of paranoid vs. nonparanoid and acute vs. chronic patients) and to isolate patterns of cognitive style-personality trait interrelations. (Study done in collaboration with J. Silverman of NIMH and D. N. Jackson of University of Western Ontario.)

Longitudinal Study of Creativity (NICHHD, 782-5) N. Kogan

PURPOSE: The long-term stability of measures of creative ability has not been systematically explored. Fifth-grade children who participated in the Pankove-Kogan study (Journal of Personality, 1968, 36, 420-439) have been retested after an interval of five years. Approximately 75% of the original sample was still in the same school systems and hence accessible for retesting. In addition to the readministration of the Wallach-Kogan tasks, the subjects were given a brief questionnaire intended to tap extracurricular accomplishments.

PROGRESS: Analysis of data has been completed, and a preliminary report was presented at the first symposium of the International Society for the Study of Behavioral Development in Nijmegen, The Netherlands, July 1971. A report (RB-71-57) has been issued and published (Creative Ability over a Five-Year Span, Child Development, 1972, 43, 427-442). A seven year follow-up upon high school graduation has been completed. Analysis of the data is currently in progress.

MAJOR FINDINGS: For boys, correlations between fifth- and tenth-grade composite creativity scores (fluency and uniqueness) were highly significant. For girls, the corresponding correlations were negligible. On the other hand, neither fifth- nor tenth-grade creativity scores were related to extracurricular activities in boys, whereas tenth-grade creativity was marginally associated with such activity in girls. In the latter case, test anxiety operated as a moderator, low anxious girls manifesting a strong positive correlation between tenth-grade fluency and extracurricular activities, in contrast to high anxious girls for whom the correlation was essentially zero.

Creativity, Learning, and Attention (NICHHD, 782-26) W. Ward

PURPOSE: To develop and validate a nonverbal creativity test suitable for use with preschool children.

PROGRESS: A task has been developed in which children are given a supply of variously colored forms and asked to construct a familiar object, for example, a tree. The task has been included in the test batteries of several large studies of young children which are in progress at ETS.

MAJOR FINDINGS: The task engages the attention of young children for prolonged periods of time; sometimes evokes highly original constructions; and can be scored reliably and inexpensively for appropriateness, unusualness, and flexibility. Analyses in progress will provide information relevant to construct validation of these scores.

Cognition of Affect (NICHHD, 783-2) S. Messick

PURPOSE: This study investigates individual differences in the conceptual differentiation of affects as a function of cognitive styles and ideological commitment.

PROGRESS: Several measures of cognition and personality have been administered to over 600 high school students along with a task in which they were asked to judge the degree of similarity or difference among 18 affect names, such as fear, surprise, joy, and contempt. These judgments of similarity were analyzed by multidimensional scaling procedures to isolate consistent individual viewpoints about affect interrelations. The first two dimensions of judged similarity clearly reflect positive vs. negative affect, and intensity (calm vs. excited). The third dimension is tentatively interpreted as reflective of negative feelings about the self (ashamed, humiliated, afraid) vs. negative feelings directed externally (contemptuous, mad, angry). This last dimension may correspond to intropunitiveness vs. extrapunitiveness. The number of stimulus dimensions and other structural properties of each viewpoint will be used to assess consistent individual differences in the conceptual differentiation of affects. The content of the stimulus representation for each viewpoint will be used to characterize the cognition of affect according to that viewpoint. Individual differences in conceptions of affect will be studied in relation to the cognitive and personality measures included in the battery, but in particular to measures of skill in discriminating and identifying visually displayed affects, the spontaneous tendency to characterize photographs of faces in affect terms, choice of preferred affect, and ideological polarity. The scoring of these procedures is currently in progress. (Study done in collaboration with S. S. Tomkins, City University of New York.)

Dimensions and Categories of Physiognomic Judgment (NICHHD, 784-2) N. Kogan

PURPOSE: The general purpose of the present study is to examine the manner in which the use of an age judgment scale varies with the chronological age of the judge.

PROGRESS: Photos of stimulus persons varying in age were presented to samples of college students and elderly people for judgments of chronological age. Both male and female stimulus persons and subjects were employed. Analyses of variance were carried out separately for male and female photos. A report based on this research is in preparation.

MAJOR FINDINGS: A main effect for judge's age was obtained only in the case of the male photos--older judges attributing higher age levels to the stimulus persons. For both male and female photos, a highly significant interaction (age of judge by age of stimulus person) was obtained. Stimulus persons at the younger end of the age continuum were judged consistently older by the older judges. The inflection point is approximately age 40 for males and age 30 for females. Stimulus persons exceeding those ages are judged differently depending upon whether they are male or female. Photos of older females are judged consistently younger by the elderly subjects, whereas the age of the judge does not seem to have a systematic effect on the ages attributed to the photos of older males.

Determinants of Risk Taking for Others
(NICHHD, 784-4)

N. Kogan
A. Teger

PURPOSE: The present research represents an extension of earlier work by Zaleska and Kogan (RB-69-9) on the comparison of individual and group decision-making for other people. The study examines the effects of friendship and reciprocity of decision-making between self and other on preferred risk levels.

PROGRESS: A report (RB-71-56) has been issued and submitted for publication. A condensed version was published in the Proceedings of the 79th Annual Convention, APA, 1971, 247-248, and a more extended version has been submitted to a journal for publication.

MAJOR FINDINGS: Decision-making for others was not affected by extent of friendship between the decision-maker and the other. On the other hand, a significant reciprocity effect was obtained. In deciding for another, subjects were inclined to take greater risk under nonreciprocal conditions than in the case where the other would in turn be deciding for the decision-maker herself.

Motivation and Performance in Adult-Evaluative and Peer-Competitive Contexts (NICHHD, 784-4)

N. Kogan

PURPOSE: In the last 15 years there has been a growing body of research into motivational factors in achievement in both the classroom and other situations. Such research has been possible because of the development of useful measures of achievement motivation. Although psychologists have long recognized the influence of the interpersonal context upon the individual's striving and performance, there has been little research on the specific effects of such contexts. The principal aim of this study is to examine the impact of two commonly occurring contexts--adult-evaluative and peer-competitive--on the achievement motivation and performance of elementary school boys.

PROGRESS: A report (RB-71-49) has been issued, and a version for publication is currently in preparation.

MAJOR FINDINGS: On the whole, the results in both the peer-competitive and adult-evaluative contexts conformed to Atkinson's motivational model of risk taking--i.e., a preference for intermediate risk taking in "hope for success" subjects and a preference for more extreme risk or caution in "fear of failure" subjects. Of further interest was evidence showing significant variation in risk levels across contexts for high defensive boys--more cautious under "adult evaluation" and more risky under "peer competition"--and the absence of such variation for low defensive boys. (Project conducted in collaboration with Mr. J. Damm of the University of Queensland, Australia.)

Cognitive Styles and Attitudinal Consistency (NICHHD, 784-9) — S. Messick

PURPOSE: Some of the major concerns of social psychology in recent years have dealt with the dynamics of attitudinal consistency--in particular, (a) with the extent to which both cognitive and affective components of attitude systems tend to be organized in psychologically consistent patterns, (b) with the operation of motivational pressures toward consistency, and (c) with the specific mechanisms whereby inconsistency is resolved. The present study investigates the importance of reliable individual differences in the tendency toward attitudinal consistency, the generality of their operation across different measurement approaches and different response processes (such as interpersonal perception and opinion change), and the extent of their relation to other cognitive and personality variables.

PROGRESS: Attitudinal consistency was measured using four experimental procedures that had previously formed the basis for four major studies in this area: The McGuire Logical Consistency Test, the Morrissette Balance Completion Test, the Osgood and Tannenbaum Congruity Test, and the Rosenberg and Abelson Fenwick Test.

These four experimental procedures were administered, along with a measure of intolerance of attitudinal inconsistency, to a sample of 200 high school seniors who had been extensively tested earlier with a battery of cognitive and personality measures. One ancillary function of the present study will be an evaluation of the replicability with high school students of the mean effects observed in the four previous experiments by Morrissette, McGuire, Rosenberg and Abelson, and Tannenbaum. The major emphasis, however, is upon the consistency of interrelations among these four procedures, upon the number and nature of dimensions required to account for individual differences in attitudinal consistency, and upon the cognitive and personality correlates of those dimensions. Particular attention will be paid to the relations between categorizing styles and attitudinal consistency, since level of conceptual differentiation and cognitive complexity may moderate consistency pressures. (Study done in collaboration with J. E. Singer of State University of New York at Stony Brook.)

Social Stratification (NICHHD, 784-16) L. Stricker

PURPOSE: A vast amount of research has been devoted to the identification and development of objective indexes of social stratification, such as education and occupation. Although this research has been based almost entirely on samples of whites, these indexes have often been used to measure the stratification of Negroes, despite suggestions that these indexes are not comparable for the two groups.

An inherent problem with these indexes, even when they are restricted to whites, concerns the identification of the cutting points for differentiating between the social classes. Empirical attempts to identify these points have largely been unsuccessful and a theoretical controversy continues over whether the social classes are distinct entities, reflecting discontinuities in stratification, or are arbitrary classifications that have been imposed on a continuum.

This study has two aims: (a) to identify the dimensions of social stratification for both Negroes and whites, and the indexes that best reflect these dimensions, and (b) to determine whether distinct social classes exist and, if so, whether they can be empirically differentiated.

PROGRESS: A comprehensive, highly structured interview schedule was prepared, focused on information that directly reflects the major aspects of social stratification or are important correlates of it. The schedule was used in household interviews with representative samples of Negro and white household heads in a typical northern city. The data are being analyzed.

Effects of Models on Creative Performance in Relation to Anxiety, Ability, and Sex (NICHHD, 786-10)

N. Frederiksen
F. Evans

PURPOSE: This study involves an attempt to improve performance on the Formulating Hypotheses test (which requires subjects to develop hypotheses to account for data) by presenting models of "acceptable" performance that stress either quantity or quality of responses. Results of a previous study (Journal of Educational Psychology, 1969, 60, 465-475) that involved only the "quantity" models showed that (1) the use of models increased the number of hypotheses formulated; (2) the training did not transfer to a test of divergent production (Guilford's Consequences test); (3) test anxiety was curvilinearly related to performance; and (4) there was a significant interaction of anxiety and verbal ability. The purposes of the present study were to attempt to replicate the previous findings and also to investigate (1) the effects of models emphasizing quality as well as the quantity models, (2) sex differences, and (3) the relationships of ideational fluency and verbal ability to performance.

PROGRESS: Data have been obtained from about 400 college freshmen in two eastern state colleges. New methods of scoring have been developed to yield scores reflecting quality as well as quantity. A multivariate analysis of variance has been completed and a report is available (RB-72-54).

MAJOR FINDINGS: Both the quantity and quality models were effective in modifying behavior in the expected direction. Ideational fluency was related to the number of hypotheses, and verbal ability was related to scores reflecting quality. Females were on the average superior to males with respect to scores reflecting quantity of responses. Test anxiety was not significantly associated with performance. Weak evidence of treatment-anxiety and sex-vocabulary interactions were found.

Self, Other, and Fear: The Reaction of Infants to People (NICHHD, 786-21)

M. Lewis
J. Brooks

PURPOSE: The investigation of infants' responses to strange has not been very well examined. Particularly, almost no work has been done on infants' responses to social events: different categories of people.

PROGRESS AND MAJOR FINDINGS: Infants between eight and 18 months of age were introduced to five different social events: strange adult male and female, strange four-year-old female, mother, and self. The infants' responses indicated that approach affects stimulus differentiation (in terms of fear), while age affects level of response intensity. Strangeness of the social event was not sufficient to explain the results, and ethological and cognitive theories were discussed. It was suggested that the cognitive construct of self be introduced to explain the results. It appears reasonable to assume that by eight months the self as differentiated from other is well established. Moreover, we propose that some specific categories of self, such as size, sex, or efficacy, may be established as well. The paper was presented at the Eastern Psychological Association and is available as RB-72-23. A more complete version will be presented at the Infant and Early Childhood Education Conference, ETS, October 1973 and will be published by John Wiley & Sons (M. Lewis, Editor).

Developmental Studies of Children's Incidental Learning and Distractibility (NICHHD, 786-23) — G. Hale

PURPOSE: This research looks at children's performance on various measures of attention, with an emphasis on identifying developmental changes in children's tendency to exercise selective attention.

PROGRESS: During the present fiscal year, Research Bulletins were issued reporting three studies conducted on this project: (1) "The Effect of Pictorial Integration on Children's Incidental Learning" by G. A. Hale and R. A. Piper (RB-73-26); (2) "Age Differences in Children's Performance on Measures of Component Selection and Incidental Learning" by G. A. Hale and S. S. Taweel (RB-73-36); (3) "Dimension Preference and Component Selection: Alternative Measures of Children's Attention to Stimulus Components" by G. A. Hale and L. E. Lipps (RB-73-48). An overview of these and earlier studies on the project is presented in a theoretical paper, "The Development of Attention in Children" by J. W. Hagen and G. A. Hale (RB-73-16), to appear in A. D. Pick (Ed.), Minnesota Symposium on Child Psychology, Vol. 7, University of Minnesota Press.

MAJOR FINDINGS: In the study by Hale and Taweel, children's attention to the nondominant component of stimuli was found to increase from ages 5 to 8 when this feature could serve as a functional cue but not when it was incidental to the purpose of the task. This result suggests a developmental increase in the flexibility of attention deployment, as children's tendency to exercise selective attention varies with the task requirements. Hale and Piper found that children's incidental learning of pictorial stimulus elements was higher when those elements were presented in a scene with the task-relevant information than when they were presented separately. However, pictorial integration produced little change in the developmental pattern of results; this developmental pattern--no increase with age in incidental learning along with a marked increase in intentional learning--points to a developmental increase in children's ability to attend only to task-relevant information. The theoretical paper by Hagen and Hale stresses that, as children grow older, they become more efficient in deployment of attention, exercising selective attention when it is advantageous to do so, but broadening their scope of attention under other circumstances. In the Hale and Lipps study, two tasks assumed to measure children's attention to stimulus attributes produced conflicting developmental results with children from ages 3 to 6. Apparently, age differences in the way children respond in the "dimension preference" test--a commonly used stimulus-matching task--do not reflect developmental changes in the way children naturally deploy attention in learning, as measured by a component selection task.

Difficulty Factors in Learning to Spell (NICHHD, 786-26) — L. Cahen

PURPOSE: The purpose of this research is to identify stimulus properties of spelling words that are associated with the spelling difficulty of the words.

PROGRESS: The literature on spelling difficulties has been reviewed and published as Research Bulletin 69-72. A modified version of this report has been published in the Review of Educational Research. A data bank which will be utilized to identify stimulus properties of spelling words that are associated with spelling difficulty of the words has been built. Data (proportion of students passing a given spelling word) from the New Iowa Spelling Scale will form the dependent variables in the analyses. A set of linguistic variables has been constructed for the predictor matrix. A Research Bulletin (70-47) has been issued. This report provides a mathematical model for the prediction of spelling difficulty. A draft of a second Research Bulletin has been prepared, and this reports a more sophisticated regression model using interaction terms. A shortened version of this paper was presented at the 1971 meeting of the American Educational Research Association. A first set of tables of phoneme-grapheme and grapheme-phoneme correspondences has been developed. The orthographic system of Professor Richard Venezky (University of Wisconsin) has been used and the frequencies have been weighted by the data from the Brown University Study (Kučera-Francis, principal investigators). This first set of tables has been weighted by types. It is planned that the correspondences will also be weighted by tokens. A comparison of the prediction utility of these two different types of tables will be compared. Both sets of correspondences have been broken down for the initial, medial, and final position in a word. A special study in predicting the difficulty of compound words is under way.

In a second study, an experiment was run which will add information about the effect of the semantic usage of words on spelling difficulty. The experiment utilized fourth- and fifth-grade students. Two forms of the spelling test were created and administered to matched groups of students. One form contained a word in one context with the same word in another context being assigned to the second form. For example, the word listed was given in one spelling test in the following format: "The boy listed the words on his paper." The second form utilized the following sentence: "The boat listed from side to side." The spelling tests were given orally and the data are now being transcribed from tape to coding sheets and eventually will be on punched cards for analyses. Differences in proportion of students passing the two words in the different contexts will be studied to determine the relative contribution of spelling ability, the general frequency of the word in the language, and parameters of context as determined from the Lorge semantic count list on spelling performance.

Developmental Differences in Stimulus Encoding (NICHHD, 786-28) — W. Ward

PURPOSE: This series of studies is directed toward an understanding of developmental changes in encoding and representational processes. It is proposed that simplistic notions such as Bruner's, that preoperational children can store information only ikonically, are wrong; rather, young children may simply be less able or less skilled at adapting their encoding strategy to task demands.

PROGRESS: Access to information through recognition memory appears to be unaffected by organizational and retrieval strategies which are crucial in tasks involving recall. Recognition tasks, therefore, can provide unconfounded information on encoding and representational processes.

Several studies have been completed in which subjects were shown a large number of pictures to be remembered, followed by a series of yes-no or forced choice tests. Pictures were presented by slide projector; subjects responded by button-pressing, providing measures of both latencies and errors. Variations in instructions and in the dimensions along which target and distractor stimuli differ provide the basis for comparisons which support inferences as to what attributes of presented information are encoded.

MAJOR FINDINGS: One study (RB-73-13) found nursery school and college age students to be highly similar in their relative performance when tested for storage of object identity and color information; these results question the generality of claims that young children are more oriented to color information than are adults. Moreover, little effect was found for instructional variations as to the nature of the subject's activities during learning, suggesting that preferred encoding strategies are relatively resistant to alteration in both young and old subjects. A second study, investigating the modifiability of encoding strategy in adults, showed that such subjects can be led by more extreme instructional conditions to adopt primarily either a verbal or a visual encoding approach, but that both kinds of information are encoded and are employed in a recognition test.

Other data, involving latencies in forced choice vs. single choice recognition tasks, provide information useful for the development of models of the information processing routines followed by subjects in recognition tasks. These data suggest that subjects in a forced-choice task engage in several comparison operations which are in fact totally redundant.

Individual Difference Parameters in Recognition Memory W. Ward
(NICHHD, 786-28)

PURPOSE: On the basis of recognition memory data, Sternberg has proposed a model of recognition memory in which successive stages of processing involve stimulus encoding, memory search, decision as to a match, and response organization and output. The present study sought to discover whether individuals show consistency across occasions and tasks in characteristics of two of these processing stages, those dealing with encoding and with response organization. It further sought to discover how such individual differences would relate to several popular cognitive style measures.

PROGRESS: College students were given a Sternberg-type recognition memory task on one occasion, and on a second day were given either a similar task or a perceptual matching task which was highly analogous to the memory task. They were also given a cognitive style battery including measures of reflection--impulsivity, analytic style, and color-word reading speed and interference. Analysis is in progress.

MAJOR FINDINGS: Individuals show substantial consistency in the efficiency of the stimulus encoding stage of processing, as measured through the effect on response latencies of stimulus degrading across both testing occasions and type of test. Changes in the probability of positive and negative exemplars, used to index characteristics of the response organization stage, show consistent individual differences across occasions for the same type of task, but not across tasks. On the perceptual matching task, stimulus degrading and response probability effects interact, suggesting that subjects adopt different information search strategies depending on the relative frequencies of the several types of events.

A preliminary report was presented at the 1972 meetings of the Eastern Psychological Association.

Manipulated Expectancies and Teacher Bias (NICHHD, 787-28)

L. Cahen

PURPOSE: The purpose of the study is to examine teacher bias and manipulated expectancies and correlates of these behaviors.

PROGRESS: Data are being analyzed to understand further the psychological processes associated with teacher bias or influenceability as conceptualized in the study. Preliminary analyses indicate that the amount of bias contributed to manipulated expectancies has been found to be a function of the amount of information given to a teacher to establish the expectancy. A draft of a Research Bulletin dealing with this analysis has been written. A second analysis will be performed on the data that have been gathered. The second analysis will focus on the personality dimensions of teachers which are associated with the amount of bias displayed in the scoring of pupils' papers. The second analysis will be sequential in nature. It will include separate analyses at different stages of scoring a set of four tests. Normative data for the tests used in the study have been obtained under controlled conditions. These data were obtained on approximately 250 subjects.

Developmental Changes in Self/World Differentiation (NICHHD, 787-30)

A. H. Baker

According to certain leading developmental theorists (e.g., Piaget, Werner, Schachtel), boundaries between self and world are lacking in human infancy and evolve only with higher levels of development. In particular, it is assumed that the relationship between self and world can vary from one of nondifferentiation (i.e., fusion or linkage) which prevails in infancy to one of relative differentiation (i.e., separateness, distance from, or oppositeness) which characterizes the adult. Several interrelated studies have attempted to demonstrate that even among normal adults, degree of self/world differentiation can vary, and that such variation has measurable effects upon perception. We have recently extended this work (1) by studying age changes in perceptual phenomena, using children as subjects; and (2) by studying individual differences in perceptual experience among adult subjects along one or more dimensions of self/world differentiation.

Study 1

PURPOSE: In studying age changes in perceptual phenomena, the specific dependent variable chosen for study is the psychophysical error of anticipation (= undershooting) which has long been known to occur in space localization experiments when some variable stimulus, presented in an ascending or descending trial, is moved toward some standard value. The central hypothesis which shaped our earlier studies with adults--viz., that with increase in degree of self/world differentiation there is an increase in magnitude of undershooting--was derived in part from developmental theory. As such, it would seem a most natural extension of this work to explore directly whether there are changes during ontogenesis in the effects of these variables which the earlier research demonstrated affects adult performance. It is thus proposed to undertake a series of studies, cross-sectional in design, using children between the ages of 7 and 18, while at the same time extending our work with adult subjects.

PROGRESS: Progress has been achieved to date in the following areas: (1) a laboratory for conducting the planned perceptual-cognitive experiments has been built; (2) necessary equipment has been constructed; (3) pertinent literature review has been accomplished; (4) a well qualified research assistant, with some expertise in the area of perception, was hired and trained in a variety of dark-room laboratory procedures which are necessary for conducting the planned research; and (5) data from some published studies by other authors have been procured and analyses have been undertaken to clarify issues relevant to this project.

The most widely supported finding emerging from this series of interrelated studies is that there is a systematic shift from undershooting error (stopping too soon) toward overshooting error (letting the stimulus go too far) over the course of repeated trials. This shift phenomenon occurs both when the experimenter and when the subject controls the stimulus; both when the stimulus is moved continuous or in steps; both with the psychophysical

methods of limits and of adjustment. It occurred over the course of several short sessions (where level of fatigue is presumably constant). We have interpreted this phenomenon as reflecting a shift from an active (= greater arousal) to a passive (= lesser arousal) orientation due to task repetitiveness. For the types of situations employed, in which the task was externally imposed and the stimulation repetitive, it was consequently hypothesized--in the context of a test of this interpretation--that the younger the child, the quicker the tendency to become bored and thus the more rapid occurrence of the shift phenomenon. This hypothesis was recently confirmed, using children ranging from 6 to 18 years of age.

Preliminary pilot studies have raised the possibility that (1) degree of undershooting/overshooting error (independent of the shift phenomenon) and/or (2) the shift phenomenon itself may constitute individual-differences variables reflecting underlying perceptual-style and personality characteristics. To explore this possibility, an extensive battery of tests has been devised which provides multiple measures of both undershooting errors and of shift in such errors over repeated trials. This battery, coupled with many potentially fruitful personality and cognitive style measures, has been administered to 120 individuals. Data analyses are now in progress. In this context, a procedure for obtaining both types of errors in a group administered setting has been devised.

Study 2

PURPOSE: The present study has as its focus individual differences in the degree of self/world differentiation among normal adult subjects. Although a number of studies in the past decade have demonstrated that variation in the relationship between self and stimulus object does have measurable effects upon perceptual experience, for the most part such studies have been experimental in nature--that is, conditions were varied, and the goal was to demonstrate that some general effect resulted from this variation. Studies of individual differences in the ongoing relationship between self and object have been quite rare and almost always limited to a single perceptual situation. The present research seeks to go beyond this by attempting--for the first time--to explore whether there are consistent individual differences in self/world orientation which can be assessed by a battery of tests tapping a variety of perceptual and cognitive situations.

PROGRESS: Progress has been achieved to date in the following areas: (1) space has been procured and a laboratory built at the Psychology Department of Florida State University for conducting the planned perceptual-cognitive experiments; (2) all equipment has been designed and is now in operational use; (3) plans have been finalized regarding the initial phases of this project; (4) an extensive set of materials was prepared specifying in detail the procedures, instructions, ways of recording data, etc. for each of the tasks decided upon; (5) the person who will be collecting the data during the initial phases was trained in the appropriate laboratory procedures; (6) to insure that standard dark-room lab procedures would obtain, the present researcher spent two days at the Florida State University laboratory checking out all of the procedures, and (7) an initial study involving part of the

total proposed battery has been conducted. This involved the standard rod-and-frame test; a similar test in which a rod alone is presented to the subject; two variants of the rod-and-frame test--one with instructions which make a person highly aware of his own body during the task, and another in which the person is asked to suppress all information from his body while carrying out the task (these latter two based on some research by Baker, 1968); and a novel task, using the rod-and-frame apparatus, in which the subject is asked to line up the stimulus rod with the longitudinal axis of his tilted body (after Baker, 1963).

Results of some of the above findings have been or will be reported as follows:

(1) Habituation (overshooting) and anticipation (undershooting) errors in visual and tactual space localization. Presented at meetings of The Psychonomic Society, November 1969, St. Louis, Missouri.
(2) Effects of stimulus structure, mode of inspection, and repeated trials upon anticipation-habituation errors in space localization. Presented at the meetings of the American Psychological Association, September 1970, Miami, Florida (see also ETS RB 70-45).
(3) Changes in body perception during repeated trials. Presented at the meetings of the American Psychological Association, September 1970, Miami, Florida (with Jerry Schlater and Seymour Wapner).
(4) Shift from undershooting toward overshooting error over repeated trials with a body adjustment task. Presented at meetings of the Psychonomic Society, November 1970, San Antonio, Texas (with Theodore Leventhal).
(5) Psychophysical error of anticipation and the method of limits: A developmental analysis. Presented at meetings of the International Society for the Study of Behavioral Development, July 1971, Nijmegen, The Netherlands.
(6) Shift from undershooting toward overshooting error over repeated trials: A developmental analysis. To be presented at meetings of the International Society for the Study of Behavioral Development, August 1973, Ann Arbor, Michigan (with Laraine M. Schwartz and Brian Mishara).
(7) Toward formalizing the "undershooting to overshooting shift" as a general phenomenon: Findings with subject-stimulus location fixed. To be presented at the Fourteenth Annual Meeting of the Psychonomic Society, November 1973, St. Louis, Missouri (with Laraine M. Schwartz and Brian Mishara).

(Study done in collaboration with Seymour Wapner from Clark University and Joseph Grosslight of Florida State University.)

Newspaper Test (ANPAF, 788-1) P. Diederich

PURPOSE: The American Newspaper Publishers Association Foundation (ANPAF) sponsors a "Newspaper in the Classroom" program that now reaches over 5 million students in grades 7-12 per year. The Foundation sought ETS help in developing a test to measure some of the effects of this program.

PROGRESS: Two forms of a test of newspaper reading for junior high school and two for senior high school were developed by trying out 400 items in newspaper classes and control classes (without newspaper training), matched on general reading comprehension scores. Only those items on which newspaper classes surpassed control classes were retained in the published forms. The test was based on two simulated newspapers of four pages each, tabloid size, which students read during the test.

MAJOR FINDINGS: In the norming administration, the four published forms were administered to 13,000 students in and around Charlotte, N. C., Peoria, Ill., and Fort Worth, Texas. Again the newspaper classes surpassed regular classes (without newspaper instruction) by a wide and significant margin. The need for such instruction is indicated by the fact that the average score of regular classes in junior high schools was 50 per cent correct; in senior high schools, 60 per cent correct. The items were not intended to be hard or tricky. They were written and reviewed by committees of teachers with the idea that anyone who understood an article should be able to answer the questions on it. The test is published and distributed by Cooperative Tests and Services.

Factorization of Covariance Matrices in Terms of a Specified Pattern or Structure (NSF, 791-1)

W. Kristof
B. Wingersky

PURPOSE: Let $\underline{\Sigma}$ be a given covariance or correlation matrix of order n x n and $\underline{P}$ a specified matrix of order n x m and full column rank, m < n. The basic problem solved here is the determination of a symmetric nonsingular matrix $\hat{\underline{\phi}}$ of order m x m and a diagonal matrix $\hat{\underline{U}}$ of order n x n such that, in a least-squares sense, $\underline{\Sigma} = \underline{P}\,\hat{\underline{\phi}}\,\underline{P}' + \hat{\underline{U}}$. Hence we are seeking a set of uniquenesses in $\underline{U}$ and a factor variance-covariance matrix $\hat{\underline{\phi}}$ that are maximally compatible with $\underline{\Sigma}$ and $\underline{P}$.

PROGRESS: The work was presented at the Psychometric Society Meeting, 1971. A Research Bulletin is in preparation.

MAJOR FINDINGS: Solutions for $\hat{\underline{\phi}}$ and $\hat{\underline{U}}$ are obtained in closed form without the use of differentiation. A computer program has been written. The method has been applied to a problem of factor comparison in a cross-cultural setting.

Testing a Linear Relation between True Scores of Two Measures (NSF, 791-1) — W. Kristof

PURPOSE: We concern ourselves with the hypothesis that two variables have a perfect disattenuated correlation, hence measure the same trait except for errors of measurement. This hypothesis is equivalent to saying, within the adopted model, that true scores of two psychological tests satisfy a linear relation. Statistical methods for testing this hypothesis are sought.

PROGRESS: The work is contained in RB-71-63 and has been published in Psychometrika, 1973, 38, 101-111.

MAJOR FINDINGS: Statistical tests of the hypothesis of a linear relation between true scores of two psychological measures are derived when the relation is specified with the exception of the additive constant. Two approaches are presented and various assumptions concerning the error parameters are used. Then the results are reinterpreted in terms of the possible existence of an unspecified linear relation between true scores of two psychological tests. A numerical example is appended by way of illustration.

Testing Whether a Disattenuated Correlation Is Perfect (NSF, 791-1) — W. Kristof

PURPOSE: A perfect disattenuated correlation between two psychological tests indicates that the two tests measure the same trait except for errors of measurement. Statistical methods for testing this hypothesis are sought.

PROGRESS: The work was presented at the 80th Annual Convention, 1972, of APA and is contained in the Proceedings, 1972, 43-44.

MAJOR FINDINGS: Statistical tests of the hypothesis of a perfect disattenuated correlation between two psychological measures are obtained. The derivation is based on the theory contained in RB-71-63. A condensed version of the theory was presented at the 1972 APA Convention, the basic approach being a correlational one. A numerical example will be given.

International Study of Educational Attainment in French as a Foreign Language (USOE through Teachers College, Columbia University, 798-19) — J. Carroll

PURPOSE: This is a part of the International Study of Educational Attainment (I.E.A.) being conducted by an International Council headed by Professor Torsten Husén of the University of Stockholm, and in which achievement tests in English as a Foreign Language, French as a Foreign Language, Civics, Literature, Reading Comprehension, and Science were administered in a number of countries for the purpose of comparative studies. The French as a Foreign Language tests were monitored by an international committee of which Mr. Carroll is chairman and Mr. Clark is a member. Tests of listening, speaking, reading, and writing were prepared for various levels of achievement, ranging from that of a 10-year-old population to that of the pre-university level.

PROGRESS: The tests approved by the International Council of the IEA Study in December 1969, with minor modifications, were administered in a number of countries in the spring of 1971 to large, statistically controlled samples of students in grades 10 and 12. A set of tests were also administered to 10-year-old pupils in the United States. The administration of the tests was handled by national centers in the various countries.

During 1972-73, the computer analyses of the large amount of data generated by this study gradually became available. (These analyses were performed partly at Teachers College, Columbia University under the direction of Richard Wolf, and partly at the University of Stockholm under the direction of Dr. T. N. Postlethwaite, Mr. Roy Phillipps, and others, with the assistance of Dr. Carolyn Massad of ETS who was on secondment to the IEA Headquarters under a Spencer Foundation Fellowship.) The final data analyses did not become available until April 1973, but Mr. Carroll was able to give a preliminary report of between-country findings at the AERA meeting in New Orleans in February 1973, under the title, "Factors accounting for between-student, between-school, and between-nation differences in performance in French as a foreign language." One of the more striking findings was that variation in the learning attainments in the various countries was to a substantial extent a function of the amount of time spent in learning. For most of the countries and populations, the mean scores on the proficiency tests were highly correlated with the average number of years that French had been studied. Mr. Carroll plans to spend a large part of the summer of 1973 writing the final report, with the assistance of Mr. Clark and Mrs. Massad; a book based on this report is planned for publication in 1973 or 1974. Analysis of the data will examine the respective roles of home and social background factors, instructional factors, and student factors in measured achievement in French as a foreign language. The published report will parallel the books that became available in May 1973 concerning other subject-matters (International Studies in Evaluation. Stockholm: Almqvist & Wiksell, and New York: Halsted Press Division of John Wiley. I: Science Education in Nineteen Countries, by L. C. Comber and J. P. Keeves; II: Literature Education in Ten Countries, by A. C. Purves; III: Reading Comprehension Education in Fifteen Countries, by R. L. Thorndike.)

Conference on Psychophysical Measurement — M. Levine
(NSF Grant to Mathematical Social Sciences Board
of the Center for Advanced Study in Behavioral Sciences,
798-56)

PURPOSE: To consolidate recent theoretical and experimental advances in the study of direct measurement (magnitude estimation, bisection and related procedures).

PROGRESS: Preprints and letters were circulated. A three-day informal discussion of issues and problems has taken place at ETS. Participants were Norman Anderson, Douglas Carroll, David Cross, David Krantz, Michael Levine, Gregory Lockhead, R. Duncan Luce, Lawrence Marks, and Stanley Rule. Written contributions were received from Roger Shepard and Robert Teghtsoonian.

Study of College Degrees — J. Warren
(Western College Association, 801-9)

PURPOSE: As college degrees become more important as a social and occupational credential, their meaning and substance become correspondingly more critical. A series of meetings with about a dozen representatives of member colleges of the Western College Association has been held to examine issues in the meaning of degrees that would be useful topics for a detailed study.

PROGRESS: The committee sketched out a domain of potential studies, and 55 representatives of WCA institutions rated the importance of 15 issues the committee had identified. A proposal for external funding of a study to identify the range and variety of competencies associated with degrees has been submitted.

Mother-Infant Interaction
(Spencer Foundation, 815-1)

M. Lewis
S. Painter

PURPOSE: To explore the ongoing interaction between the mother and her 12-week-old infant, and to follow through and investigate this communication as the infant develops through its first and second years of life.

PROGRESS: Ninety-seven 12-week-olds of both sexes and from 5 social classes have been seen with their mothers in their own homes. The data collected include a wide variety of maternal and infant behaviors. Emphasis has been placed on observing the chains of interaction that form the communication network of the mother and her three-month-old infant. These 12-week-old infants have been brought to the laboratory by their mothers after the home visit. They have been given measures of visual and auditory attention, which have been shown to be associated with (related to) cognitive development. At the same time, they have also been given some of the standard measures of development, such as the Bayley scales. The data from this phase of the research are being analyzed and some preliminary results are discussed below.

When the same children reached the age of one year, they were again brought to the laboratory and were observed in a standardized free-play situation. Observation methods included a dictated continuous behavioral account of the infants' activities, and the recording of four measures of infant attachment and behavior. The last of the mother-infant pairs has recently been seen and preparations are being made to analyze these data. When the data have been analyzed, we hope to be able to make some definite statements about the relationship between mother-infant interaction as seen at 12 weeks and the attachment bond of the one-year-old.

Some of the infants involved in the study have turned two years and have been brought back to the laboratory for a final, two-session series of observations. The standardized free-play situation (modified to meet the needs of a two-year-old) is again employed, as well as several cognitive measures which we hope will lead us to some conclusions about the relationship of the nature of mother-infant interaction and cognitive development.

MAJOR FINDINGS: Some preliminary data from the 3-month phase of the project have been analyzed. Some of the major results have been the contribution of models by which data from this and other research can be made meaningful. Specifically, the model of a general communication matrix is one which takes into account the diverse patterns of mother-infant interaction which are the essence of their communication network. From this model we come to see that it is not only the frequency of behavior with which we must be concerned, but the chaining effect of initiation and response of these behaviors as well. Thus, for example, mothers of boys smile more than mothers of girls in terms of behavior frequency. However, when maternal smiling is looked at in terms of interaction, i.e., when it is an initiation of or a response to an infant behavior, mothers of girls smile more

than mothers of boys. The development of a Markov model for analyzing the vocalization chains of the mother and infant rises naturally from the general communication matrix model. By use of the Markov model, we can begin to detect the fine lines of communication between the mother and her infant. Several papers have reported on these and other preliminary findings: Lewis, M. State as an infant-environment interaction: An analysis of mother-infant interaction as a function of sex. Merrill-Palmer Quarterly, 1972, 18, 95-121; Lewis, M., and Freedle, R. Mother-infant dyad: The cradle of meaning. Research Bulletin 72-22. Princeton, N. J.: Educational Testing Service, 1972; Lewis, M., and Lee-Painter, S. An infant's interaction with its social world: The origin of meaning. Paper presented at the Canadian Psychological Association meetings, Montreal, June 1972; Lewis, M., & Lee-Painter, S. An interactional approach to the mother-infant dyad. To appear in M. Lewis and L. Rosenblum (Eds.), The origins of human behavior. New York: John Wiley & Sons, 1973, in press.

Assessment of Infant Socioemotional Development
(OCD, 817-1)

M. Lewis
L. Michalson

PURPOSE: To develop an instrument to assess the socioemotional development of children within the first two years of life. The primary use of this assessment will be to evaluate the development of infants involved in day care centers. It is hoped that such an instrument can be used to explore the effects of any kind of intervention program.

PROGRESS: To this end 12 infants were observed intensively in a day care center, and their behaviors in a variety of situations were recorded. Based on these observations, an assessment scale, related to the infants' ongoing experiences in the intervention program, was constructed.

Five domains were selected for study. Future investigations will be directed toward gathering normative data, examining the relationships among the domains, and establishing the developmental course of these socio-emotional variables through the first years of life. In addition to observing day care infants, a comparable group of home-care infants will be rated on the assessment instrument in an effort to evaluate the effects of the intervention program.

Peer Group Interaction in Infancy (OCD, 817-1)

J. Brooks
M. Lewis
L. Michalson

PURPOSE: This study was designed to explore infants' specific interactions with strangers of different ages. Specifically, we are interested in the infants' responses to peers and to adults and in the early peer group interaction. Exploration of peer group dynamics in preverbal children has been limited.

PROGRESS AND MAJOR FINDINGS: Eight groups of 12-month-old and eight groups of 18-month-old infants will be observed in a low-stress play situation. Four infants (two males and two females) and their mothers will be placed in a playroom for 15 minutes. Behaviors such as looking, smiling, touching, cooperative play, and aggressive play will be measured; these behaviors may be directed toward the mother, the adult strangers or the other infants. The type and amount of behavior directed to the adults and infants will be compared; sex differences will also be analyzed. Data are being collected and analyzed.

Developmental Studies of Selective Attention (NICHHD, 826-1) G. Hale

PURPOSE: This research addresses the question, to what extent do children attend selectively to a single aspect of stimulus materials and ignore other, extraneous information? To identify developmental changes in this process from early school age to preadolescence, two general approaches to the study of attention are combined: (a) assessment of learning and recall for the various attributes of stimuli and (b) examination of the distracting effects of extraneous stimulation.

PROGRESS: Five studies were conducted on this project during the current fiscal year; data collection is complete in all cases, preliminary analyses have been conducted, and major analyses are continuing.

MAJOR FINDINGS: In one study, 384 children of ages 5, 8 and 12 learned a task involving stimuli differing on two redundant features, and then were tested for recall of each feature separately. Instructions to attend to the nondominant feature increased recall for this feature at all ages; recall of the other feature was also increased at age 5 but was reduced at age 8 and at age 12. With increasing age, children apparently become more efficient at ignoring nonessential features while attending to task-relevant information. A second study with 264 5- and 8-year-olds used a similar paradigm but different materials and reached a similar conclusion--that is, with increasing age children become better able to accommodate attention-related instructions. In another study, involving 384 children at ages 5, 8 and 12, changes in the relative salience of two stimulus components caused attention to shift toward the more salient feature. The consistency of this effect across ages points to a general perceptual effect that may be unrelated to development of other cognitive abilities.

Children's attention to auditory distractors was measured by examining performance on a short-term memory task both in the presence and in the absence of extraneous auditory stimulation. In a study involving 108 5- , 8- and 11-year-olds, hearing annoying noises proved to be more distracting than hearing the names of the task stimuli (repeated in random order) for 5-year-olds, while the reverse was true for 8- and 11-year-olds. Apparently, as children grow to rely more heavily on verbal rehearsal, stimulation that specifically interferes with this process increases in distractiveness relative to other, nonspecific distractions. Another study, including 120 subjects at ages 5, 8 and 11, examined developmental changes in ability to divide attention between two tasks. While performing the memory task, a subject was required to tap the table whenever he heard the word "tap" amid the noise. This condition was preceded by a series of trials under non-distraction and produced a marked drop in performance at all ages, but after several trials even the 5-year-olds' performance improved to a level close to that observed under nondistraction. With practice, even children of early school age apparently can divide attention between auditory and visual information.

Development of Interest Measurement Instruments for Assignment of Army Officers to Appropriate Positions
(Army Research Institute, 827-1)

J. Campbell
R. Boldt
R. Flaugher
D. Rock

PURPOSE: This project is designed to provide the Department of the Army with dimensions of Army Officer interest appropriate for use with a computer-aided career counselling system, a set of instruments to measure these dimensions, and with similarity measurement and clustering techniques for grouping officers and positions.

PROGRESS: Visits have been made to several Army installations to discuss activity preferences with officers. Army job descriptions are being reviewed to secure activity statements.

System of Interactive Guidance and Information (Carnegie, 842-3) — M. Katz

PURPOSE: A computer-assisted system of interactive guidance and information is being developed to improve career decision-making by students in junior colleges.

SIGI is a guidance system which allows the student to interact with a computer in such a way as to examine his own values, obtain and use relevant information, interpret predictive data, and formulate plans. This interaction assists the student in arriving at tentative career decisions and in modifying them as he gains new insights and additional information. The decisions involve both educational and occupational options. Emphasis, however, is not merely on the content of decisions but on the process of decision-making. As the student progresses through SIGI, he learns to move freely within the structure of the system. In gaining control of the system, he progressively develops competencies and masters strategies for rational behavior in the face of uncertainty.

PROGRESS: A preliminary tryout of SIGI was conducted during the summer of 1972. A sample of students used the Introduction and Values system on the PDP-8 and dual-screen terminal and then went on with the Information, Prediction, and Planning systems on the PDP-11 and single-screen terminal. Observations were made for formative evaluation of SIGI. A description and illustration of SIGI appeared in the Summer 1972 issue of the EDUCOM Bulletin.

The entire system was then converted to the PDP-11 and single-screen terminal and used by a representative sample of the Mercer County Community College student body during the spring semester of 1973. Evaluative data were collected from control and experimental groups. Data are now being analyzed. A report will be written.

Options for various delivery systems for SIGI have been studied.

Validating Effects of SIGI on Career Decision-Making Processes of Students (NSF, 842-4) — M. Katz

PURPOSE: To observe the effects of SIGI on the quality and nature of community college students' reasoning and information for career decision-making.

PROGRESS: Criteria were formulated, and interview schedules and other measures were constructed and tried out during the summer of 1972. Revised instruments were used in conjunction with a field test of SIGI at Mercer County Community College during the spring semester of 1973. Data have now been collected and are being analyzed. A report will be written.

Departmental Characteristics and Student Academic Growth (844-1) — J. Centra

PURPOSE: (1) To identify departments that have graduates with higher or lower achievement than predicted; and (2) to identify characteristics of these most and least effective departments.

PROGRESS: This study uses the "input-output" model to investigate departmental effects on student cognitive growth in college. Student achievement on the Graduate Record Examinations Advanced Tests or the Undergraduate Program Tests (ETS) were used as the (output) measure of departmental effectiveness; student college admissions scores (SAT or ACT) or their National Merit Scholarship Qualifying Test scores are the input measures.

Development and Validity of an Air Force Vocational Inventory Study (847-1)

G. Echternacht
R. Reilly

PURPOSE: The purpose of this project is to develop and validate an Air Force Vocational Interest Inventory for use in the Air Force Guaranteed Assignment Program. A total of 12 homogeneous vocational interest scales will be developed from an initial pool of 400 items. The scales will be validated using a sample of approximately 3,200 airmen and eight selected career fields (two each of the four Airman Qualifying Examination areas).

PROGRESS: A 400-item inventory has been developed and administered to over 3000 airmen in eight selected career fields. Satisfied career groups will be defined and classification rules will be developed to distinguish between men-in-general and satisfied personnel within each career field. Both empirical and homogeneous scales will be developed and tried.

Status of Drug-Related Instructional Programs in Secondary and Elementary Educational Institutions
(Commission on Marijuana and Drug Abuse, 849-1)

R. Boldt
R. Reilly

PURPOSE: Numerous statistics may be assembled to the effect that marijuana and drug use may be fairly common in the age ranges extending from secondary school down. The purpose of the study was to explore these and related data through a survey of school principals.

PROGRESS: A review of literature reporting on the evaluation of drug abuse educational programs and the results of the survey were reported to the Commission as a partial basis for its recommendations to the President.

MAJOR FINDINGS: The findings of the Commission are available through the Government Printing Office, and the result of the present survey will be reported in full in an appendix to that report. Publication of the appendix is expected in August 1973 and an ETS Project Report (PR-73-6) will be made available 60 days later.

Assessment and Development of Cognitive Skills in Young Children: A Classroom Approach (Carnegie, 858-6)

E. Chittenden
A. Bussis

PURPOSE: The central purpose of this two-year project, initiated in July 1970, is to complete development of materials that will help teachers understand and foster the intellectual development of children in pre-kindergarten through grade three. Goals are to provide the teacher with specific guides and techniques to help observe, interpret, and encourage significant behaviors in children between approximately 4-9 years of age. The final materials will incorporate and thus supersede the current Let's Look at First Graders material.

PROGRESS: Tasks and materials for the early grade levels are nearing a final stage of completion and are based largely on previous research phases of this project which were also supported by Carnegie. During the past year, additional measures have also been developed and tried out at the 1st, 2nd, and 3rd grade levels. Major emphasis is being placed on four aspects of development: language functioning and communication skills; quantitative thinking; classification and ordering; use of egocentric and intuitive processes.

Study of Teacher Change in Open Educational Settings
(865-19)

T. Chittenden
A. Bussis
M. A. Amarel

PURPOSE: This project is a teacher interview study addressed to what are considered two major facets of educational experimentation and accountability: (1) effective implementation of an educational approach, and (2) the creation of an environment that will sustain and promote continuing growth within that approach. Initially, at least, our concern with these questions will be focused on "open" educational environments and research on the teacher's role--her educational priorities, perceptions, attitudes, working assumptions. Specific goals of the project are as follows: (1) to identify and describe concrete issues and factors which teachers perceive as both facilitating and inhibiting of change toward an open classroom; (2) to link characteristics of teachers, their environment, and their perceived needs for assistance into a better theoretical understanding of the dynamics of change; and (3) in the course of the research, to refine an interview instrument and set of scoring procedures which would be applicable to teachers working in innovative reform programs other than open education.

PROGRESS: Sixty teachers in urban public elementary schools have been interviewed. Data from these in-depth interviews are being analyzed.

Student Points of View in Ratings of College Instruction (865-20)

J. Centra

PURPOSE: The purpose of this study was to identify subgroups of students with identifiable "points of view" in their ratings or descriptions of instruction they had received. In addition to identifying dimensions that described the students, the relationship of individual characteristics, such as expected grade in the course and overall grade point average, to those dimensions would be studied.

PROGRESS: Inverted factor analyses within each of three classes indicated four to five subgroups of students. Further analyses are now under way with 300 randomly selected students (from some 400 classes).

NSF Science Aptitude Trends Study (NSF, 865-24) R. Boldt

PURPOSE: Over a period of time the National Science Foundation has contributed substantial funds to graduate education in science. In recent years there has been a well-publicized overage of technical personnel produced by higher education. The purpose of the study was to utilize historical data based on the Graduate Record Examinations which might indicate trends in aptitude in those who have taken the examinations, particularly those who have taken them in the sciences.

PROGRESS: The history files of the Graduate Record Examinations were used to construct a history of aptitude test statistics for the years 1966-67, 1967-68, 1970-71, and 1971-72. For the latter three years the classification of students by fields was accomplished using the department which they indicated should receive their scores: A change of reporting procedure dictated that for data for 1966-67, the classification should be in the field in which they indicated an intention to study.

MAJOR FINDINGS: The 1966-67 means were depressed with respect to the others, and their comparability is questioned. Declining trends in means were noted for the last three years studied for physical science, math science, engineering, basic social science, applied social science, and the arts and humanities. A declining trend in the Quantitative score was very strong for physical science, math science, and engineering, and a declining trend in Verbal scores was also noted in the other areas mentioned above. The research is reported in a Project Report (PR-73-7).

Test and Job Performance of Various Subgroups within Specified Occupations (Ford Foundation, 879-3)

J. Campbell
L. Crooks
D. Rock
M. Mahoney

PURPOSE: This project was designed to investigate the relationship between job and test performance for different ethnic groups when pertinent background factors are taken into consideration. The project was carried out under a grant from the Ford Foundation as a joint responsibility of ETS and the U. S. Civil Service Commission.

PROGRESS: The relationship of job and test performance for different ethnic groups was studied in three different occupations: medical technician, cartographic technician, and inventory manager. In each of these federal Civil Service occupations, samples of job incumbents were given a battery of aptitude tests selected after careful analysis of job duties and responsibilities. A background, experience, and job task questionnaire was also completed, and superiors rated each incumbent on specially constructed experimental rating scales. Job knowledge tests were also constructed for medical technicians and cartographic technicians and work sample tasks for cartographic technicians and inventory managers. Samples of Black and Caucasian job incumbents were available for all three jobs, and of Mexican-American for the cartographic technicians and inventory manager jobs.

MAJOR FINDINGS: Principal findings include the following:

1. Aptitude tests which have validity in relation to job performance for one ethnic group generally show validity for other ethnic groups as well.

2. Tests which are valid against a rating criterion also show validity against more objective criterion measures.

3. Multiple regression weights determined on a single ethnic group hold up surprisingly well on cross-validation across different ethnic groups.

4. Ethnic group rater-ratee combinations interact to affect the ratings assigned, but the effect appears to be complex and probably differs from one ethnic group to another.

Results of the medical technician project have been reported in Project Reports 69-5, 69-6, 70-16, 70-17, 70-18. Development of the instrumentation has been reported in Project Reports 69-4, 71-22, and 71-23. The proceedings of an invitational conference to discuss and critique the findings has been published, and the complete technical report will be published shortly.

The Effectiveness of Student Feedback in Modifying College Instruction (894-1)

J. Centra

PURPOSE: The effects of formal student feedback to college instructors were investigated in this study through three related questions: Do instructors learn something about their teaching from this feedback? Do instructors improve (i.e., do student ratings improve) after receiving the feedback? And which particular instructional procedures or teacher behaviors tend to be changed?

PROGRESS: Using a specially developed instrument to obtain student reactions to instruction, student feedback was given half way through a semester to one group of instructors (the experimental group) and withheld from a second group (the control group). The instrument was readministered to both groups at the end of the semester as well as to a post hoc group, and comparisons in changes were made. The number of years of teaching experience, sex of the instructor, instructor self-ratings, and subject area of the course were also examined for their relationship to changes.

Data were collected at five diverse colleges during the Fall 1971 semester. Additional ratings were collected at one of the colleges at the end of the Spring 1972 semester.

MAJOR FINDINGS: Changes in instruction occurred after only a half semester for instructors whose self-evaluations were considerably better than their student ratings. Variables such as subject area of the course, sex of the instructor, and number of years the instructor had taught did not distinguish which instructors made changes. However, a wider variety of teachers changed if given more than a half semester of time and if they received some minimal information to help them interpret their ratings. Results are reported in SIR Report #2 (PR-72-16).

Self-Ratings of College Teachers: A Comparison to Student Ratings (894-1) — J. Centra

PURPOSE: The purpose of this study was to investigate teachers' self-ratings of instruction by comparing them to ratings given by students. Discrepancies between self-ratings (or self-descriptions) and those provided by students would underscore the need for student feedback as well as highlight specific areas of instruction where feedback is most essential.

PROGRESS: The sample consisted of 343 teaching faculty from five colleges; these teachers, as well as the students in one of their classes, responded to a 21-item instructional report questionnaire. The data were collected during the mid-semester of Fall 1972, and analyses have been completed.

MAJOR FINDINGS: Correlating teacher responses to each item with the mean class responses (across the 343 classes) disclosed a modest relationship between the two sets of evaluation: a median correlation of .21 for the items. In addition to the general lack of agreement between self- and student evaluations, there was also a tendency for teachers as a group to give themselves better ratings than their students did. Comparisons between student and faculty responses were also made across items, and a rank correlation of .77 indicated a good deal of similarity in the way the two groups rank ordered the items.

Discrepancies between individual teacher ratings and ratings given by the class were further analyzed for (a) sex of the teacher (no difference found); (b) number of years of teaching experience (no difference); and (c) subject area of the course (differences noted for natural science courses vs. those in education and applied areas). Among other conclusions, the results of this study would argue for the collection of student ratings to supplement self-ratings. The results are reported in RB-72-33 and as part of SIR Report #2.

Bell Lab Day Care Evaluation (895-2) V. Shipma

PURPOSE: The Bell Lab Day Care Project, begun last year, was designed to evaluate an experimental day care center established by AT&T. The project has been funded by Bell Labs, acting in an advisory capacity to AT&T.

PROGRESS: As in the past year, the evaluation included use of a parent interview, individual child tests, teacher questionnaire and ratings, and classroom observations to provide information on parent, child, and teache behaviors and attitudes towards the center. A final report will be issued this fall following completion of data analyses now in process.